GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Elliptical Extended Object Tracking

Dissertation
for the award of the degree
"Doctor rerum naturalium" (Dr. rer. nat.)
of the Georg-August-Universität Göttingen

within the doctoral program Ph.D. Programme in Computer Science (PCS)
of the Georg-August University School of Science (GAUSS)

submitted by
Shishan Yang
from Shandong, China P. R.

Göttingen, 2021

Thesis Advisory Committee

Prof. Dr.-Ing. Marcus Baum,
Institute of Computer Science, Georg-August-Universität Göttingen, Germany

Prof. Dr. Dieter Hogrefe,
Institute of Computer Science, Georg-August-Universität Göttingen, Germany

Members of the Examination Board

Referee:	Prof. Dr.-Ing. Marcus Baum, Institute of Computer Science, Georg-August-Universität Göttingen, Germany
Co-referee:	Prof. Simon Godsill, Department of Engineering, University of Cambridge, United Kingdom
Second Co-referee:	Asst. Prof. Florian Meyer, Scripps Institution of Oceanography and Electrical and Computer Engineering Department, University of California San Diego, United States

Further Members of the Examination Board

Prof. Dr. Dieter Hogrefe,
Institute of Computer Science, Georg-August-Universität Göttingen, Germany

Prof. Dr. Carsten Damm,
Institute of Computer Science, Georg-August-Universität Göttingen, Germany

Prof. Dr. Xiaoming Fu,
Institute of Computer Science, Georg-August-Universität Göttingen, Germany

Date of the Oral Examination
04.12.2020

I hereby declare that I have written this thesis independently without any help from others and without the use of documents or aids other than those stated. I have mentioned all used sources and cited them correctly according to established academic citation rules.

Göttingen, 17.05.2021

Acknowledgment

I would like to thank the following people for supporting me in completing this thesis.

First and foremost, I would like to express my sincere gratitude to Marcus Baum for guiding me into such a fascinating research area and supporting me when I worked at the University of Göttingen and the University of Passau. Furthermore, I am grateful to professor Dieter Hogrefe for being my second supervisor and advising me in the past years.

I was thrilled when Marcus told me that professor Simon Godsill and professor Florian Meyer agreed to be the co-referees of this thesis. Their pioneer work on Poisson Point Process modeling and data association for extended objects inspired me and influenced me enormously. In addition, my gratitude also goes to professor Damm and professor Fu for being the members of my examination committee.

I am so lucky to work in the Data Fusion lab and have such great colleagues. Fabian Sigges, Kolja Thormann, and Laura Wolf, thank you for all the inspiring discussions in the past years. Besides, my appreciation goes to the friends that I have made since I started in Göttingen: Shradha Fowdur, Florian Teich, Hauke Kaulbersch. It gives me immense pleasure to thank my colleges in the Institute of Computer Science. Especially thanks go to my colleague but also my dearest friend Hang Zhang for picking up my daughter Angela when I was occupied.

In the end, I would like to thank my parents and parents-in-law for their love and support. Thank you all for coming to Germany and helping me take care of my daughter to support me in finishing this thesis. I would like to thank my husband, Chuan Zhao, for believing in me even when I am not. Last but not least, Angela, thank you for coming into my life and motivating me to be a person that you will be proud of.

Göttingen, October 2020 Shishan Yang

Abstract

Extended object tracking is an emerging research topic that is motivated by the rapid development of modern sensors. The traditional object tracking assumes a tracked object is far away from the sensor. Therefore, an object takes only one resolution cell and can be simplified as a point. However, due to the employment of near-field and high-resolution sensors, it is common for an object to occupy several resolution cells, and its extent is not negligible in many modern applications such as autonomous driving, robotics, and surveillance. Extended object tracking estimates both the kinematic state and spatial extension of an object based on a varying and unknown number of measurements. In this thesis, the object extensions are described as elliptical shapes. This thesis is devoted to three problems in the context of extended object tracking and has made three contributions respectively:

Evaluation metric

Between two ellipses that describe the same object, which one is better? Many elliptical extended object trackers have been developed, but no consensus exists on the measures for performance comparison. The Euclidean distance, which evaluates the location error for point object trackers, incorporates no shape error. Finding a simple and intuitive measure that combines both location and shape errors is not straightforward. Through the discussion and evaluation of the possible performance measures, the first contribution of this thesis is the proposal of using the Gaussian Wasserstein distance for evaluating elliptical extended object trackers.

Shape estimation

Given a set of measurements originated from one extended object, how to derive the kinematics and the shape of the underlying object? The estimation of object extension is challenging as it is a high-dimensional and non-linear estimation problem. The state-of-the-art elliptical trackers approximate the object shape as a symmetric positive definite random matrix, which couples the

orientation and axes lengths. However, modeling the dynamics of orientation and axes lengths individually is useful for many applications. Therefore, the second contribution of this thesis is a single elliptical extended object tracker that explicitly estimates object kinematic state, orientation and semi-axes lengths. A closed-form solution is derived in the framework of recursive Kalman filter. Using the Gaussian Wasserstein distance as a metric, simulation results have shown that the proposed tracker facilitates the dynamic modeling of extended objects and outperforms the previous work on this topic.

Multiple extended object tracking

Knowing a set of measurements from multiple objects, what are the location and the shape of each object? The key to solve this problem is data association, i.e., determining the origin of each measurement. Many multiple extended object trackers rely on clustering techniques to obtain measurement partitions so that measurements generated from the same object are in one cell. Then, the measurement cells are assigned to potential objects using data association methods in traditional object tracking. However, the clustering process normally incorporates predicted object density heuristically and has high complexity. The third contribution of this thesis is a new multiple extended object tracker that employs an efficient measurement-object assignment method and using the single extended object tracker for shape estimation in contribution two. The new data association method calculates the marginal association probabilities by considering all measurement-object mappings, yet requires no clustering or explicit enumeration of assignments. The proposed tracker is tested using simulation and real lidar data. Results showed that the proposed tracker is more efficient and performs better than clustering-based trackers.

Contents

1 Introduction **1**

1.1 Motivation . . . 1

1.2 Extended Object Modeling . . . 4

1.2.1 Shape Representation . . . 4

1.2.2 Measurement Modeling . . . 5

1.2.3 Dynamics Modeling . . . 6

1.3 Considered Problems and Challenges . . . 6

1.3.1 Metrics for Evaluation . . . 6

1.3.2 Shape Estimation . . . 7

1.3.3 Multiple Extended Object Tracking . . . 7

1.4 Contributions . . . 7

1.5 Structure of this Thesis . . . 8

2 Metrics for Extended Object Trackers Comparison **9**

2.1 Problem Description . . . 10

2.2 Metrics for Multiple Object Tracking . . . 12

2.3 Measures for Elliptical Shapes Comparison . . . 14

2.3.1 L_p Distance . . . 14

2.3.2 Decoupled Measures . . . 15

2.3.3 Intersection over Union . . . 16

2.4 Closed-Form Measures for Gaussian Representation . . . 19

2.4.1 Kullback-Leibler Divergence . . . 19

2.4.2 Hellinger Distance . . . 20

2.4.3 Gaussian Wasserstein Distance . . . 20

2.4.4 Comparison on Closded-Form Probabilistic Measures . . . 21
2.5 Uniform Wasserstein Distance by Sampling . . . 22
2.6 Conclusion . . . 27

3 Shape Tracking with Explicit Parameterization 29
3.1 Related Work . . . 30
3.2 Recursive Bayesian Tracking . . . 31
3.2.1 Kalman Filter . . . 32
3.2.2 Extended Kalman Filter . . . 33
3.3 Problem Description . . . 34
3.4 Modeling an Elliptical Extended Object . . . 34
3.4.1 Parameterization . . . 35
3.4.2 Measurement Equation . . . 35
3.4.3 Dynamic Model . . . 37
3.5 Estimation based on a Tailored Extended Kalman Filter . . . 39
3.5.1 Sequential Measurement Update . . . 40
3.5.2 Time Update . . . 44
3.5.3 Derivations . . . 45
3.6 Evaluation . . . 48
3.6.1 Moment Approximation . . . 48
3.6.2 Constant Velocity Model . . . 50
3.6.3 Coordinated Turn Model . . . 50
3.7 Discussion . . . 53
3.7.1 Relationship with Random Matrix Approaches . . . 53
3.7.2 Parameter Selection . . . 56
3.8 Conclusion . . . 57

4 Multiple Extended Object Tracking without Measurement Partitioning 61
4.1 Related Work . . . 63
4.1.1 Multiple Point Object Tracking . . . 63
4.1.2 Multiple Extended Object Tracking . . . 65
4.2 Preliminaries . . . 66
4.2.1 Probabilistic Data Association Filter . . . 66
4.2.2 Joint Probabilistic Data Association . . . 68
4.2.3 Joint Integrated Probabilistic Data Association . . . 69
4.3 Problem Description . . . 72
4.3.1 Assumptions and Approximations . . . 74
4.3.2 Joint Association Hypothesis for Extended Objects . . . 74
4.3.3 Zero-Inflated Poisson Point Process . . . 74
4.3.4 Existence Probability . . . 75

4.3.5 Predicted Measurement Likelihood 75
4.4 State Prediction 76
4.5 Data Association 77
4.5.1 The Posterior Probability of a Joint Association Hypothesis 77
4.5.2 The Marginal Association Probability 80
4.5.3 Existence Probability Update 82
4.5.4 Conditional Marginal Association Probability 85
4.6 State Update 85
4.7 Implementation 86
4.7.1 Birth Model and Predicted Likelihoods 87
4.7.2 Data Association 89
4.7.3 Track Management and Weighted MEM-EKF* 90
4.7.4 Discussion 90
4.8 Evaluation 92
4.8.1 An Intuition on the Proposed Data Association Method 92
4.8.2 Simulation with Closely-Spaced Objects 96
4.8.3 Tracking Pedestrians using Lidar Data 99
4.9 Conclusion 103

5 Conclusion **105**

Acronyms **107**

Nomenclature **109**

Own Publications **113**

List of Figures

1.1 An Example shows radar measurements from nuScence dataset 2
1.2 An Example shows the lidar data and corresponding image from KITTI dataset . . 2
1.3 Examples of point and extended object tracking 3
1.4 An example of the description complexity of the object shape 4
1.5 Two measurement source models of an extended object 5

2.1 An example of evaluating single elliptical extended object trackers 11
2.2 An example of evaluating multiple extended object trackers 12
2.3 An example showing the calculation of OSPA distance for point objects 13
2.4 A counter-intuitive example of using L_p distance for ellipses comparison 15
2.5 Example of using decouple measures for elliptical shape evaluation 17
2.6 An example showing IoU lacks a closed-form solution for parameterized elliptical shapes 18
2.7 A counter-intuitive example of using IoU for extended objects comparison 18
2.8 A counter-intuitive example of using the Kullback-Leibler divergence and the Hellinger distance to compare ellipses 21
2.9 The counter-intuitive example of the Kullback-Leibler divergence and the Hellinger distance 23
2.10 An example that shows the different ordering given by Gaussian Wasserstein distance and a uniform Wasserstein distance approximated by sampling on boundaries 25
2.11 Different associations of uniform Wasserstein distances for different numbers of samples 25
2.12 The constructed scenario for comparing Gaussian Wasserstein and Uniform Wasserstein distances 26

3.1 The derivation of the measurement equation using multiplicative noise 36
3.2 Simulation with a static ellipse to evaluate the moments approximation 49
3.3 The measurements, trajectory, and estimation results of a single example run . . . 51
3.4 Extension and location error based on the root mean squared Gaussian Wasserstein distance 52
3.5 The trajectory and estimates of the coordinated turn simulation 54
3.6 Root mean squared Gaussian Wasserstein distance for 100 runs 55
3.7 The effect of different shape covariance matrices in the prior 58

4.1 The two-step data association method for multiple extended object tracking 62
4.2 An illustrative example of the proposed data association method by calculating the marginal association probabilities 63
4.3 A PDA example 66
4.4 A JPDA example for point object tracking 69
4.5 Markov chain one model for object existence 71
4.6 The structure of MEM-JIPDA filter 73
4.7 The possible existence and detection statuses of object τ 76
4.8 An example illustrating the proposed marginal association probability and existence probability for extended object tracking 93
4.9 Simulated sensitiveness test of marginal association probabilities in terms of Poisson rates 95
4.10 Simulated sensitiveness test of existence probabilities in terms of Poisson rates. . . 95
4.11 The simulation setup and one example run of the GGIW-PHD filter and the MEM-JIPDA filter 97
4.12 The comparison of GGIW-PHD and MEM-JIPDA using mean OSPA distance $\pm$ one standard deviation for the simulation with closely-spaced objects 99
4.13 Pedestrian Lidar measurements in the KITTI campus dataset 100
4.14 The pedestrian tracking result of MEM-JIPDA using the lidar data from the KITTI dataset 102

List of Tables

3.1 The ground truth and prior for evaluating moment approximation 48
3.2 The setting of parameters in the constant velocity simulation 50

4.1 The marginalized association probabilities obtained for the example in Figure 4.8 . 94

1 Introduction

Contents

1.1 Motivation 1
1.2 Extended Object Modeling 4
1.3 Considered Problems and Challenges 6
1.4 Contributions 7
1.5 Structure of this Thesis 8

Object tracking is ubiquitous in our daily life and humans track various kinds of objects unconsciously. This makes object tracking one crucial task in many computer-aided applications such as robotics, surveillance and autonomous driving. Object tracking is a broad area and has many different settings. This thesis is constrained to the classic non-cooperative object tracking. To be more specific, it focuses on the process of recursively estimating and predicting the states of dynamic objects based on a series of measurements that is originated from a remote sensor. A typical example is aircraft surveillance (see Figure 1.3(a)).

1.1 Motivation

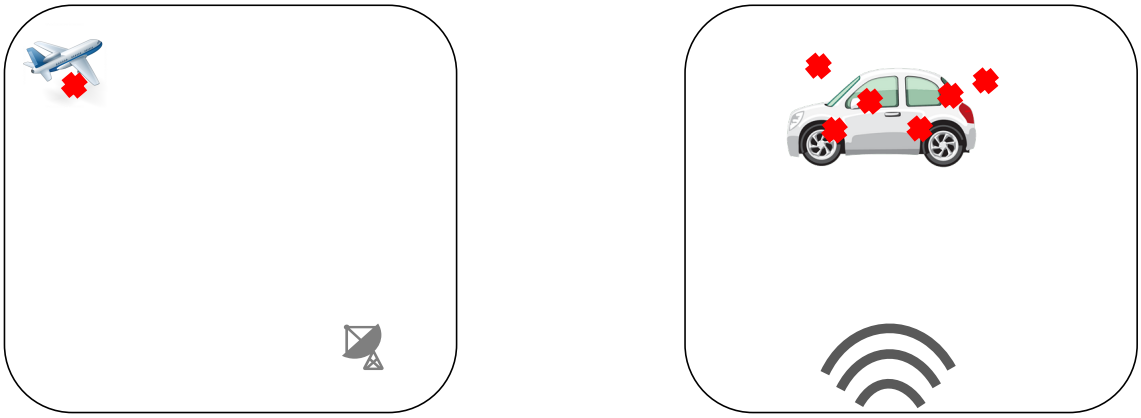
The traditional object tracking assumes the tracked object is far away from the sensor and has the so-called *small object* assumption. Under the small object assumption, the extent of the object is neglected and the object is treated as a point. In addition, it assumes one object generates at most one measurement.

With the rapid advances of sensor technology, the small object assumption is not valid for many applications. Figure 1.1 gives an example of the measurements received from an automotive radar, which are mounted on a vehicle [1].[1] Figure 1.2 shows the image and reflection points that are

[1] https://www.nuscenes.org/

Figure 1.1: This figure shows the radar measurements from nuScenes dataset [1]. The left image is the bird view of the scene. The ego vehicle is indicated using a patched blue rectangle, radar measurements are plotted by red points, and the annotated vehicles are orange rectangles. The right figure overlays the radar measurements on the camera image with colors indicating depth.

Figure 1.2: Example image and lidar data from KITTI dataset [2]. The top row is the images data and the bottom row are the corresponding lidar data with two different views. As the field-of-view of camera and lidar are different, not all objects and pedestrians in the lidar data can be found on the image.



(a) Point object tracking (b) Extended object tracking

Figure 1.3: Examples of point and extended object tracking. As shown in (a), the object in point object tracking is far away from the sensor and assumes one object generates at most one measurements. Hence, the object extent can be treated as a point. (b) depicts a typical scenario of extended object tracking. For extended object trakcing, the tracked object generates a varying number of measurements, most likely more than one. Besides the location of the object, the extension of the object is one of the key estimation tasks for extended object trackers.

obtained from an RGB camera and a lidar [2].[2] From both figures, we can see that the pedestrians and vehicles occupy more than one resolution cell of the corresponding sensor. The extension of the object cannot be neglected in applications such as advanced driving assistance systems for safety considerations.

The so-called research topic *Extended Object Tracking (EOT)* is becoming increasingly important. In contrast to traditional point object tracking, the small object assumption is not valid. In addition to kinematics of the object, estimating the extension of an object is one crucial task of EOT. [3] Besides, EOT assumes that one object can generate a varying number of measurements. Figure 1.3 gives an example showing the differences between traditional object tracking and EOT.

This thesis differentiates the concepts of EOT, group object tracking, and object tracking with multi-path propagation. All of these three tracking tasks assume one tracking entity has the possibility of generating multiple measurements. However, a group object tracking method tracks a set of objects that share some common motion properties while the dynamics of an individual object can still deviate from others. For example, a group object tracker is commonly required to have the mechanism to deal with group splitting and merging, which is not necessary for EOT. The multi-path propagation tracking mainly addresses the multi-detection effects as in over-the-horizon radars. The signals from an over-the-horizon radar can be scattered and arrive at the

[2] http://www.cvlibs.net/datasets/kitti/

[3] The extended object in this thesis is considered as a combination of the kinematics of center and its spatial extent. The words shape, extent, and extension all refer to the spatial extension confined by orientation and size, with respect to the center. Therefore, shape, extent, and extension are used interchangeably.

(a) (b) (c)

Figure 1.4: An example of three levels of description complexity of the shape of an object, i.e., a pedestrian in this case, in two-dimensional space. The shape of the pedestrian is neglected in (a), approximated using an ellipse in (b), and described by an arbitrary shape in (c). In general, a higher level of description is more computational complex and requires more information from the measurements, i.e., more measurements and less sensor noise.

receiver due to the transmission in the ionospheric layers. Normally, there is a maximum number of possible propagation paths and hence a maximum number of measurements for one object. Besides, the multi-path propagation tracking normally does not estimate the object extension and assumes no spatial model for measurements.

1.2 Extended Object Modeling

Extended object tracking is a broad research topic and has different assumptions depending on the application domains. This section explains the scope of the extended object tracking that is considered in this thesis.

1.2.1 Shape Representation

In [3, 4], the shapes of objects are categorized according to the dimension and the description complexity of the shapes. For example, in a two-dimensional space, a point object is considered as zero-dimensional, a line or curve is one dimensional, and an ellipse is two-dimensional. Depending on the levels of description complexity, the shape of an object can be neglected, simplified as basic geometric shapes such as sticks, rectangles, and ellipses, or represented as arbitrary shapes. An example of difference description complexity in the two-dimensional space is given in Figure 1.4. In addition to computational cost, the description level of shape depends on the measurements. With a higher level of description, the measurements are expected to be denser and less noisy. A geometric shape provides some shape information yet can be defined using a number of parameters.

Remark 1.1. Considering the characteristics of the measurements provided by automotive radar and lidar, this thesis focuses on approximating the object extensions as elliptical shapes (see Figure 1.4(b)). A two-dimensional elliptical shape can be represented using a fixed number of

(a) (b)

Figure 1.5: Two measurement source models of an extended object. The measurements are plotted using red crosses and the measurement sources are green dots. The measurement model in (a) assumes the measurement sources on the boundary while the model in (b) describes measurement sources lying on the surface. The arrows indicate sensor noise in both sub-figures.

parameters, which balances the level of shape description and computational complexity. Elliptical shape approximation is one of the most common extension representations in the state-of-the-art studies [5, 6, 7, 8, 9].

1.2.2 Measurement Modeling

In EOT tasks, a measurement has two layers of uncertainty. First, same as in the traditional point object tracking, measurements are corrupted with sensor noises. In addition, the origin of a measurement, which is called a measurement source, is distributed differently depending on many factors, such as the type of the sensor, perspective of the sensor, the distance between the sensor and the tracked object, and the shape of the object. A measurement model should incorporate both types of uncertainty. Same as in the classic object tracking, we assume the sensor noises are additive white Gaussian noises. Depending on what types of sensors are used, the measurement sources can be approximated as reflection points that lie on the surface [10, 11] or the boundary [12, 13]. Examples are given in Figure 1.5.

Furthermore, a measurement model of EOT should interpret the number of measurements generated from an object. As the EOT problem originates from point object tracking, many of the early studies model an extended object as a set of fixed reflection points on a rigid body [14, 15, 16]. Like a point object, each reflection point generates at most one measurement and are detected independently. Therefore, the number of measurements from an extended object cannot exceed the number of reflection points. This measurement model simplifies an EOT problem as a traditional multi-object tracking problem. This kind of measurement model is categorized as *a set of points on a rigid body* in [4] and *exact/approximate rigid-body* in [17].

Another kind of measurement model assumes that the number of measurements from one object follows a probabilistic distribution. One of the most popular measurement models falling in this

category is the spatial Poisson Point Process (PPP) model proposed in [18, 19]. It models the measurements from an object as an inhomogeneous PPP. Each measurement from the extended object is spatially distributed with a probability density proportional to the intensity of the PPP. Therefore, this model is called the *PPP spatial model* in [4], or the *approximate Poisson-body* in [17]. The spatial PPP model has been one of the most popular measurement models in EOT and widely used in different literuature [11, 20, 5, 21]. Other probabilistic models exist. For example, studies in [22, 23] assume the number of measurements is Binomial distributed.

Remark 1.2. The measurement model employed in this thesis follows the spatial PPP model in [18, 19]. To be more specific, this thesis considers the EOT problem by assuming

- the number of measurements is Poisson distributed,
- the measurement noises are additive white Gaussian noises, and
- the measurement sources are distributed on the surface of the object, which is approximated by an ellipse.

1.2.3 Dynamics Modeling

In this thesis, the dynamic model of an extended object is the same as in the point object tracking and is described as a Markov process, i.e., the current object state depends only on the state from the previous time step. Section 3.2, 3.4.3, and 4.4 give more detailed explanation on dynamic models that are considered in this thesis.

1.3 Considered Problems and Challenges

EOT is an emerging research topic that is investigated in recently decades. It has various promising problems to be investigated. This thesis focuses on three sub-problems in EOT.

1.3.1 Metrics for Evaluation

Given an extended object tracker, one of the crucial questions is: *how good is it?* Given several extended object trackers, we want to know which one is the best. The performance evaluation becomes increasingly important as the development of various EOT methods. For evaluation, we aim at a score that measures the distance between an estimate and a ground truth. As EOT estimates the kinematic and shape information of an object, the estimation error on both aspects should be considered in the evaluation. The classic object tracking methods use the Euclidean distance as a standard metric for evaluation. However, the Euclidean distance has many shortcomings when it comes to the evaluation of EOT algorithms. Many state-of-the-art studies decouple the estimation error as location error, kinematics error, shape error, and other criteria. Decoupled distances give more intuition of the performance, but a single score is favored in many cases. Therefore, the first

question investigated in this thesis is: *what distance is most suitable for comparing ellipses in the context of EOT?* We aim at a measure that tells us the precision of a tracker, follows the intuition of humans, and is consistent. In addition, we also expect it to be a true metric, to have fewer parameters, and be efficient to compute.

1.3.2 Shape Estimation

As it is mentioned above, one of the most important aspects that make EOT different from classic object tracking is that the extension needs to be estimated given a varying number of measurements. The second question considered in this thesis is: *what is the kinematics and shape of an object given a set of measurements that originate from the underlying object?* Estimating the kinematic and shape of an object is challenging as it is a high-dimensional estimation problem and non-linearity lies in the measurement modeling. The standard filtering approach for point objects cannot directly be applied without the adoption of tailored approximations and assumptions. Since multiple measurements are received in one scan, information on the object extension can be extracted from the spatial distribution of the measurements. Normally, the entire shape is not observed in one scan as the measurements are often noisy and sparse. However, information from previous scans can be incorporated using filtering techniques.

1.3.3 Multiple Extended Object Tracking

For most tracking applications, multiple objects are present in the tracking area and the measurement origins are unknown. This problem is called *Multiple Object Tracking (MOT)*. The goal of MOT is to estimate the time-varying states of a number of objects in the presence of noisy measurements, clutter, and miss-detections. The third problem explored in this thesis is *multiple extended object tracking with unknown object-measurement associations*. MOT is a challenging task even for point objects. One intuitive approach for MOT is to decompose it as a multiple single object tracking tasks. To achieve this, one needs to find the origins of the measurements and update the object state using the corresponding measurement(s). The process of determining the origins of the measurements, i.e., associating measurements and objects, is called data association in MOT. In general, regardless of the object types, data association faces the intractable complexity due to a large number of association hypotheses. In contrast to point objects, the data association problem for Multiple Extended Object Tracking (MEOT) requires more computation effort. This is because MEOT requires to grouping measurements into several measurement cells, so that one cell contains all measurements of the same object.

1.4 Contributions

The published journal and conference papers are numbered with prefix *Yang*, such as [Yang1] and [Yang2]. In conclusion, three contributions are made corresponding to the three sub-problems of

EOT explained in Section 1.3:

- *proposal on using the Gaussian Wasserstein distance for comparing elliptical objects*. There is no consistently used metric for comparing elliptical objects in existing EOT studies yet. Based on the discussion and comparison of possible measures, this thesis suggests the Gaussian Wasserstein distance for evaluating the performance of single elliptical extended object trackers. The Gaussian Wasserstein distance evaluates error on location and shape estimation, is a true metric, and has a closed-form solution for ellipses comparison.
- *a single extended object tracker which explicitly tracks orientation and axes lengths*. Shape estimation is one of the vital tasks of EOT. Many state-of-the-art elliptical extended object trackers represent the object shape as a Symmetric Positive Definite (SPD) matrix, which couples the orientation and size. The second contribution of this thesis is an extended object tracker that estimates orientation and axes lengths based on the extended Kalman filter. Explicitly tracking orientation and axes lengths facilitates the dynamic modeling of the individual parameter, which is highly relevant to many practical applications.
- *a MEOT tracker which requires no clustering or measurement enumeration*. Due to the characteristics of the data association problem for extended objects, existing multiple extended object trackers perform measurement-clustering as a pre-process to reduce the computational complexity. The third contribution of this thesis is an efficient multiple extended object tracker, which avoids clustering or measurement enumeration by the computation of marginal association probabilities.

1.5 Structure of this Thesis

This thesis consists of five chapters. This chapter outlines the background of the research topic, considered problems and contributions. Chapter 2 discusses the metrics and similarity scores that can be used for extended object tracker evaluation and suggests the Gaussian Wasserstein distance for elliptical objects comparison. Chapter 3 focuses on measurement modeling of one extended object and estimating the kinematics and shape of a single extended object given a set of measurements that is originated from the underlying object. An extended object tracker based on a multiplicative noise model and a tailored extended Kalman Filter (KF) is proposed. Chapter 4 investigates the MEOT problem without knowing the origins of the measurements. A new data association method that requires no clustering or exhaustion of measurement partitions is presented. Combining with the single EOT tracker in Chapter 3, an efficient MEOT tracker is proposed. In the end, Chapter 5 concludes this thesis and discusses future work.

2

Metrics for Extended Object Trackers Comparison

Contents

2.1 Problem Description 10
2.2 Metrics for Multiple Object Tracking 12
2.3 Measures for Elliptical Shapes Comparison 14
2.4 Closed-Form Measures for Gaussian Representation 19
2.5 Uniform Wasserstein Distance by Sampling 22
2.6 Conclusion 27

Many extended object trackers approximate the object shape using rigid shapes, such as rectangles[12, 24], ellipses [25], star-convex shapes [26, 27]. Due to the convenience of modeling, ellipses have been one of the most popular shape approximations for extended object tracking algorithms. Many elliptical extended object trackers are proposed, such as random matrices approaches [10, 5, 28], Random Hypersurface Model (RHM) [11][Yang4], independent axes estimation [29].

There are no consistently used metrics or measures for evaluating extended object trackers yet. The Euclidean distance between an estimate and ground truth is used as an evaluation metric for point objects [30, 31, 32]. An extended object tracker simultaneously tracks the object location and shape. Correspondingly, besides the location error, the object shape error also needs to be incorporated in the performance evaluation. Unfortunately, finding a score that combines both location and shape errors is not straightforward. Some computer vision tasks, such as image segmentation and object detection, are related to extended object tracking to some extent as both aim to find spatial occupancies of particular objects or classes. However, computer vision tasks work with discrete pixels while extended objects deal with parameterized shapes.

This chapter first formulates the problem of evaluating EOT trackers. Then, the Optimal Sub-Pattern Assignment (OSPA) distance, which is the de-facto standard metric for evaluating multiple

point object trackers, is explained. To adapt the OSPA distance for extended object trackers, one requires a distance for comparing ellipses. The possible measures and similarity scores for elliptical shape comparison are overviewed. The advantages and disadvantages of existing measures that are used in the EOT literature are discussed. As the elliptical shapes can be represented using Gaussian distributions, the closed-form probabilistic measures, such as the Kullback-Leibler divergence, the Wasserstein distance, and the Hellinger distance, are introduced. Hypothetical scenarios are constructed and show the counter-intuitive results of the Kullback-Leibler divergence and the Hellinger distance for elliptical EOT evaluation. In the end, uniform Wasserstein distances are explored due to the restriction of the Gaussian representation of ellipses. Based on the discussion and designed simulations, the Gaussian Wasserstein distance is suggested for assessing the performance of elliptical extended object trackers.

Remark 2.1. This chapter is based on the conference publication [Yang9], which discussed the evaluation scores for extended object tracking. Afterward, the Gaussian Wasserstein distance has been used as a metric for elliptical shapes comparison [33, 34, 35, 36, 37]. This chapter mainly focused on the measure discussion for elliptical EOT trackers. The discussion extends [Yang9] by analysing Hellinger distance (Section 2.4.2). Counter-intuitive examples of Kullback-Leibler divergence and Hellinger distance are given in Section 2.4.4. In the end, the combination of Gaussian Wasserstein distance and OSPA for MEOT tracker comparison is explained.

2.1 Problem Description

Elliptical extended object trackers recursively estimate the kinematic state and shape variable of objects based on a set of measurements. The evaluation of extended objects consists of evaluation single extended object trackers and multiple extended object trackers, see Figure 2.1 and Figure 2.2 respectively. This section first presents the formulation of single-extended-object trackers comparison. Then, the problem of evaluating multiple extended object trackers is explained.

For the comparison of single-extended-object trackers, the error between the ground truth object and an estimate is computed. We focus on one scan and consider only the position and shape variable, i.e., high order and other kinematics such as velocity and acceleration are not included in the comparison. At time instant k, the true state of an object is denoted as $\boldsymbol{x}_k \in \mathcal{S}$, where $\mathcal{S}$ is the set of possible shapes. The estimates of the unknown object is differentiated using a hat sign, i.e., $\hat{\boldsymbol{x}}_k^{(1)}$ and $\hat{\boldsymbol{x}}_k^{(2)}$ are the two estimates from tracker 1 and tracker 2 (see Figure 2.1). The goal is to find a function

$$d : \mathcal{S} \times \mathcal{S} \to \mathbb{R}^+ \tag{2.1}$$

such that $d(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k)$ gives us the similarity or distance between the ground truth object $\boldsymbol{x}_k$ and an estimate $\hat{\boldsymbol{x}}_k$. Then, the distances $d\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(1)}\right)$ and $d\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(2)}\right)$ give the quantitative evaluation of two estimates.

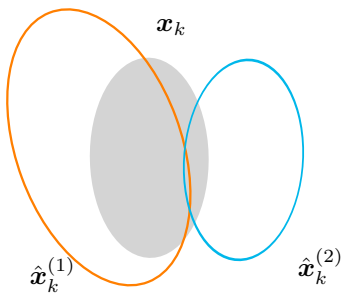


Figure 2.1: This figure shows an example task of evaluating single extended object trackers. The ground truth object is plotted by a filled ellipse. The estimates from single extended object trackers are plotted using orange and cyan ellipses. The orange estimate overlaps more with the ground truth object. But both size and orientation of the cyan estimate are similar to the ground truth object. Which ellipse is a better estimate of the filled ellipse?

Except for capturing the difference between an estimate and ground truth object, the function d is expected to have some other properties. First, it should be expressive. The "distance" function should incorporate the evaluation of the most important aspects of a tracker. Besides, the measure should be simple and fast to compute. Furthermore, the measure should be clear, easy to understand, and reflect human intuition. Ideally, d is a metric and has physical meaning. For all $\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(1)}, \hat{\boldsymbol{x}}_k^{(2)} \in \mathcal{S}$, a function d is a metric if it satisfies the following axioms:

- identity of indiscernibles: $d(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k) = 0$ if and only if $\boldsymbol{x}_k = \hat{\boldsymbol{x}}_k$,
- symmetry: $d(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k) = d(\hat{\boldsymbol{x}}_k, \boldsymbol{x}_k)$,
- triangle inequality: $d\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(1)}\right) \leq d\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(2)}\right) + d\left(\hat{\boldsymbol{x}}_k^{(2)}, \hat{\boldsymbol{x}}_k^{(1)}\right)$.

A multiple extended object tracker estimates not only the states of the object but also the number of objects presented in the surveillance area. As the number of tracked objects varies over time, the metric for MEOT evaluation needs to incorporate center error, shape error, and cardinality error. Figure 2.2 illustrates an example of evaluating multiple extended object trackers. A cardinality error can be caused by miss detected existing objects or falsely estimated non-existent objects (see Figure 2.2). Many studies have proposed measures that can be used for multiple (point) object tracking, such as OSPA, generalized OSPA. The measures that are used for traditional object tracking can be easily adapted for multiple extended object tracking given a distance function d, which is used for single EOT evaluation. Section 2.2 gives more details on the metrics of multiple extended object tracking.

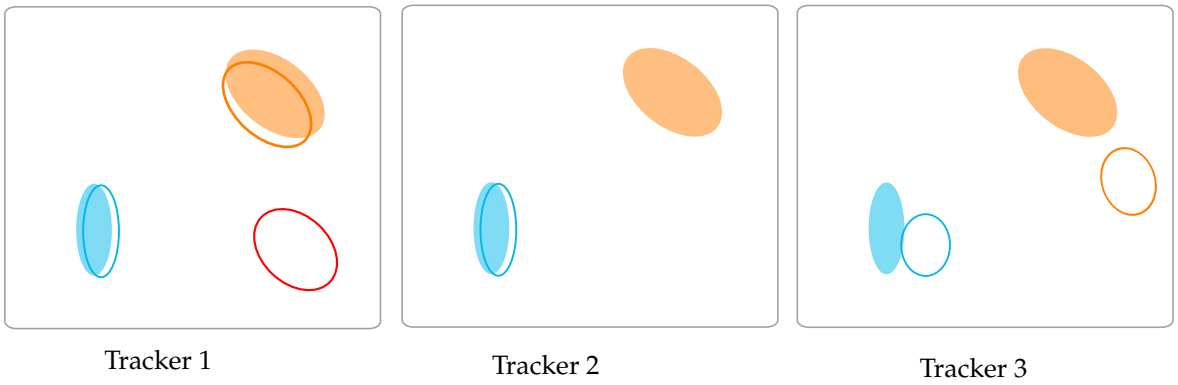


Figure 2.2: This figure illustrates an evaluation example of the multiple extended object trackers. Two objects, which are depicted by filled orange and cyan shapes, need to be tracked at a certain time instant. Three exemplar tracking results are depicted. The tracker 1 tracks both objects well but gives a false positive. The second tracker fails in tracking the orange object. The third tracker tracks both objects but the location and shape estimates are worse than tracker 1 and 2. Which tracker is the best?

2.2 Metrics for Multiple Object Tracking

The traditional object tracking algorithms treat the tracked object as a point. The Root Mean Squared Error (RMSE) between the state vectors of ground truth object and estimates is the standard metric for evaluating the single object trackers.

As the multiple object tracking involves cardinality error, various studies investigated many measures for evaluating multiple point object trackers. The CLEAR MOT measures are the de-facto standard for MOT evaluation in computer vision [38]. The CLEAR MOT measures are a set of measures that cover the accuracy, precision, and completeness of a tracker. The accuracy is indicated using MOT Accuracy, which considers the number of false positives, false negatives, and identity switches. The precision of a tracker is described by MOT Precision while the number of most tracked, partially tracked, and most lost trajectories reflect the completeness quality of a tracker. CLEAR MOT counts the cardinality errors heuristically. The Hausdorff metric and Wasserstein distance were first introduced to evaluate MOT filters in [39]. However, the Hausdorff metric is insensitive to cardinality errors. The Wasserstein distance, which is also known as the optimal mass transfer metric, is inconsistent and counter-intuitive for some scenarios with different cardinality errors. A so-called OSPA metric combines the accuracy and precision for evaluating MOT filters [30]. The OSPA distance is a true metric. It solves the problems of using Hausdorff and optimal mass transfer distance for evaluating MOT filters.

The OSPA distance is proposed in [30] and is currently the de-facto metric for evaluating multiple point object trackers. Given two sets of objects, the OSPA distance permutes the assignments from the objects in the smaller set to the other set and calculates the respective cost based on a given

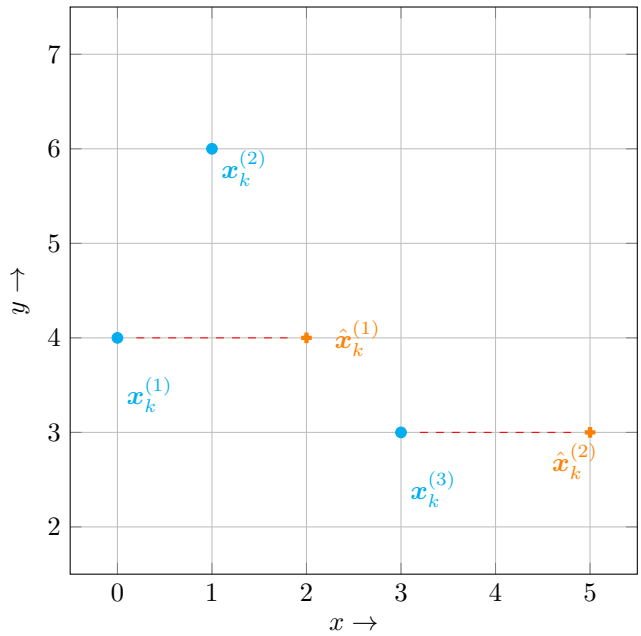


Figure 2.3: An example showing the calculation of OSPA distance for point objects. The ground truth objects are indicated using cyan dots, while the estimates are orange dots. The optimal assignment, which gives the least OSPA distance, is plotted using dashed lines.

distance function. The not-assigned objects are penalized by a so-called cut-off value to account for the cardinality error. Then, OSPA returns the smallest cost among all permutations.

At a certain time instant, given the set of ground truth $\mathbf{X}_k = \left\{\boldsymbol{x}_k^{(1)}, \cdots, \boldsymbol{x}_k^{(m)}\right\}$ and the set of estimated states is $\hat{\mathbf{X}}_k = \left\{\hat{\boldsymbol{x}}_k^{(1)}, \cdots, \hat{\boldsymbol{x}}_k^{(n)}\right\}$. Let Π_n be the sets of all permutations on $\{1, \cdots, n\}$, $d^c\left(\boldsymbol{x}_k^{(i)}, \hat{\boldsymbol{x}}_k^{(\pi_i)}\right) = \min\left(c, d\left(\boldsymbol{x}_k^{(i)}, \hat{\boldsymbol{x}}_k^{(\pi_i)}\right)\right)$ with d being a distance measure. Let the cut-off value c penalize the cardinality error between ground truth states and estimates. The OSPA distance with cut-off value c and p-th order is defined as

$$\bar{d}_p^c(\mathbf{X}_k, \hat{\mathbf{X}}_k) = \left(\frac{1}{n}\left(\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^c\left(\boldsymbol{x}_k^{(i)}, \hat{\boldsymbol{x}}_k^{(\pi_i)}\right)^p + c^p(n-m)\right)\right)^{\frac{1}{p}}, \tag{2.2}$$

if $n \geqslant m$, and $\bar{d}_p^c(\mathbf{X}_k, \hat{\mathbf{X}}_k) = \bar{d}_p^c(\hat{\mathbf{X}}_k, \mathbf{X}_k)$ if $n < m$.

In the implementation, an optimal assignment method, such as the Hungarian method [40], can be used to decide the permutation with minimum cost. Many variations of OSPA have been proposed, such as OSPA-T for including track identities [41], and Generalized OSPA for incorporating the cost of identity switches [42].

Example 2.1. Given a set of two-dimensional ground truth objects $\mathbf{X}_k = \left\{\boldsymbol{x}_k^{(1)}, \boldsymbol{x}_k^{(2)}, \boldsymbol{x}_k^{(3)}\right\}$ and the set of estimated objects $\hat{\mathbf{X}}_k = \left\{\hat{\boldsymbol{x}}_k^{(1)}, \hat{\boldsymbol{x}}_k^{(2)}\right\}$, where $\boldsymbol{x}_k^{(1)} = \begin{bmatrix}0, & 4\end{bmatrix}^{\mathrm{T}}$, $\boldsymbol{x}_k^{(2)} = \begin{bmatrix}1, & 6\end{bmatrix}$, $\boldsymbol{x}_k^{(3)} = \begin{bmatrix}3, & 3\end{bmatrix}^{\mathrm{T}}$, $\hat{\boldsymbol{x}}_k^{(1)} = \begin{bmatrix}2, & 4\end{bmatrix}^{\mathrm{T}}$, and $\hat{\boldsymbol{x}}_k^{(2)} = \begin{bmatrix}5, & 3\end{bmatrix}$. The ground truth objects and estimates are shown in Figure 2.3. There are six permutations on $\{1, 2, 3\}$. The permutation that returns the minimum distance with order of two and cut-off value of 10 is $\pi = \{1, 3, 2\}$ and $d^{10}\left(\hat{\boldsymbol{x}}_k^{(1)}, \boldsymbol{x}_k^{(1)}\right)^2 + d^{10}\left(\hat{\boldsymbol{x}}_k^{(2)}, \boldsymbol{x}_k^{(3)}\right)^2 = 8$. Adding up the penalization of cardinality errors, the OSPA distance is $\left(\frac{1}{3}\left(8 + 10^2\right)\right)^{\frac{1}{2}} = 6$.

2.3 Measures for Elliptical Shapes Comparison

In this section, the possible measures for elliptical extended object tracker assessment are presented. The possible metrics are categorized as (i) L_p distance, (ii) decoupled measures, (iii) Intersection over Union, and (iv) probabilistic measures. In this section, we discuss the advantages and disadvantages of employing L_p distance, decoupled measures, and Intersection over Union for comparing extended object trackers. The probabilistic measures will be explained in next section.

2.3.1 L_p Distance

Given extended objects that are presented using rigid shapes, i.e., a set of parameters, the L_p distance of state variable is one of the most natural metrics for comparing shapes. For example, if an elliptical object is parameterized using center $\boldsymbol{m}_k = \begin{bmatrix}m_{k,1}, & m_{k,2}\end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^2$, orientation α_k, and semi-axes lengths $l_{k,1}, l_{k,2} \in \mathbb{R}^+$, i.e.,

$$\mathcal{P}_{\boldsymbol{x}_k} = \begin{bmatrix}m_{k,1} & , m_{k,2}, & \alpha_k, & l_{k,1}, & l_{k,2}\end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^5 \ . \tag{2.3}$$

The estimate $\hat{\boldsymbol{x}}_k$ has the same parameterization. The L_p distance between $\boldsymbol{x}_k$ and $\hat{\boldsymbol{x}}_k$ is calculated element-wise for each parameter in (2.3), i.e.,

$$d_{L_p}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k) = \left(\sum_{i=1}^{5} \parallel \mathcal{P}_{\boldsymbol{x}_k,i} - \mathcal{P}_{\hat{\boldsymbol{x}}_k,i} \parallel^p\right)^{\frac{1}{p}} \ , \tag{2.4}$$

where the subscript i is the i-th element in vector $\mathcal{P}_{\boldsymbol{x}_k}$ and $\mathcal{P}_{\hat{\boldsymbol{x}}_k}$, respectively. The L_p distance is easy and a straightforward adaption of metrics from traditional object tracking. It gives us a single score, and other kinematics variables such as velocity, turn rate can be incorporated.

However, using the L_p distance for elliptical EOT evaluation has many shortcomings. In general, parametrization is not unique. Many parameter vectors can define the same shape. For example,

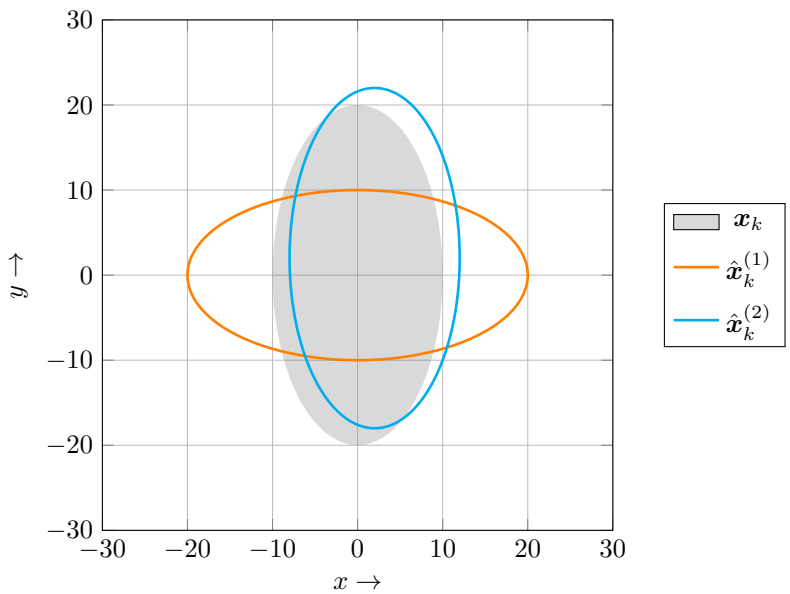


Figure 2.4: A counter-intuitive example of using L_p distance for ellipses comparison. Intuitively, the cyan ellipse is a better approximation of the filled ellipse. Using the parameterization in (2.4) and confine the orientation in the range of $\left[0, \quad \frac{\pi}{2}\right)$, the L_p distance suggests the oranges estimate has smaller error than cyan ellipse.

using the parameterization in (2.3), the filled ellipse in Figure 2.4 can be represented by many vectors, such as $\begin{bmatrix}0, & 0, & 0, & 10, & 20\end{bmatrix}^{\mathrm{T}}$ and $\begin{bmatrix}0, & 0, & \frac{\pi}{2}, & 20, & 10\end{bmatrix}^{\mathrm{T}}$. However, the L_p distance between these two vectors is not zero. The identity property in shape space is lost by directly comparing L_p distance between shape vectors. Even though identity can be assured by constraining the range of the parameters, the units of parameters are often not the same. For example, the orientation unit could be rad while the unit of semi-axes length and center could be km or m. It not only raises the problem of deciding the L_p distance unit but also makes the comparison arguable.

A counter-intuitive example is given in Figure 2.4. The ground truth object is $\mathcal{P}_{\boldsymbol{x}_k} = \begin{bmatrix}0, & 0, & 0, & 10, & 20\end{bmatrix}^{\mathrm{T}}$. The orange estimate is $\mathcal{P}_{\hat{\boldsymbol{x}}_k^{(1)}} = \begin{bmatrix}0, & 0, & \frac{\pi}{2}, & 10, & 20\end{bmatrix}^{\mathrm{T}}$ and the cyan estimate is $\mathcal{P}_{\hat{\boldsymbol{x}}_k^{(2)}} = \begin{bmatrix}2, & 2, & \frac{\pi}{2}, & 10, & 20\end{bmatrix}^{\mathrm{T}}$. As shown in the figure, typically, the cyan estimate is expected to be a better estimate for the ground truth object. However, we have L_2 distances, i.e., the Euclidean distances, as $d_{L_2}\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(1)}\right) = 40.4233$ and $d_{L_2}\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(2)}\right) = 41.0935$, which is counter-intuitive. The L_2 distance suggests that the orange ellipse is better than the cyan ellipse.

2.3.2 Decoupled Measures

To tackle the problems arisen by directly comparing L_p distance, many studies employed the decoupled errors [43, 5, 12, 10, 44]. The estimation error is decoupled as a set of errors on specific

properties, such as center error, velocity error, length and width errors, orientation error, shape error. Most of the decoupled errors are calculated using the L_p distance between the respective parameters except the shape error, which is indicated using the matrix norm and combines orientation, width, and length errors.

The shape of an ellipse can be represented as a SPD matrix. As the orientation and axes lengths are coupled in the SPD matrix representation, many studies use matrix norms to indicate the shape error, which counts for size and orientation error. Frobenius norm is used for evaluating elliptical objects in [45, 5, 46, 44].

Given a ground truth object with orientation α_k, semi-axes lengths $l_{k,1}$ and $l_{k,2}$, its shape can be represented using matrix

$$\boldsymbol{\Sigma}_{\boldsymbol{x}_k} = \begin{bmatrix} \cos\alpha_k & -\sin\alpha_k \\ \sin\alpha_k & \cos\alpha_k \end{bmatrix} \begin{bmatrix} l_{k,1}^2 & 0 \\ 0 & l_{k,2}^2 \end{bmatrix} \begin{bmatrix} \cos\alpha_k & -\sin\alpha_k \\ \sin\alpha_k & \cos\alpha_k \end{bmatrix}^{\mathrm{T}} . \tag{2.5}$$

Assume an estimate $\hat{\boldsymbol{x}}_k$ having shape matrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}$, the shape error using Frobenius norm, also known as the Euclidean distance or squared root distance [47], is

$$||\boldsymbol{\Sigma}_{\boldsymbol{x}_k} - \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}|\,|_{\mathrm{F}} = \sqrt{\mathrm{tr}\left(\left(\boldsymbol{\Sigma}_{\boldsymbol{x}_k} - \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}\right)\left(\boldsymbol{\Sigma}_{\boldsymbol{x}_k} - \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}\right)^{\mathrm{T}}\right)} . \tag{2.6}$$

The usage of decoupled distances for elliptical tracker evaluation is depicted in Figure 2.5. The ground truth object is $\mathcal{P}_{\boldsymbol{x}_k} = \begin{bmatrix} 0, & 0, & \frac{\pi}{4}, & 2, & 1 \end{bmatrix}^{\mathrm{T}}$. The orange estimate is $\mathcal{P}_{\hat{\boldsymbol{x}}_k^{(1)}} = \begin{bmatrix} 1, & 0, & \frac{\pi}{4}, & 2, & 1 \end{bmatrix}^{\mathrm{T}}$ and the cyan estimate is $\mathcal{P}_{\hat{\boldsymbol{x}}_k^{(2)}} = \begin{bmatrix} 0, & 0, & \frac{3\pi}{4}, & 3, & 1 \end{bmatrix}^{\mathrm{T}}$. The center errors of the orange and cyan estimate are $1\ \mathrm{m}$ and $0\ \mathrm{m}$, respectively. The Frobenius norm between ground truth and two estimates, which combines size and orientation, are $||\boldsymbol{\Sigma}_{\boldsymbol{x}_k} - \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k^{(1)}}||_{\mathrm{F}} = 0$ and $||\boldsymbol{\Sigma}_{\boldsymbol{x}_k} - \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k^{(2)}}||_{\mathrm{F}} = 8.5440$. Based on the decoupled measures, the orange estimate is better at shape but worse at location estimation. Quantitatively, it is not easy to decide which estimate is better overall. Decoupled distances give more information on a specific aspect of the evaluated trackers. However, it is not intuitive to find a fair weighting schema to give a single score for the overall comparison.

2.3.3 Intersection over Union

A score that counts for shape dissimilarity and center error is the Intersection over Union (IoU), which is also known as the Jaccard similarity coefficient and Jaccard index. The IoU score is widely used for measuring similarity in many computer vision tasks, such as object detection, images segmentation, and tracking[48, 49, 50, 51, 52][Yang12]. In the context of extended object tracking, it was also employed in evaluation in [12]. As the name suggests, IoU is calculated using the

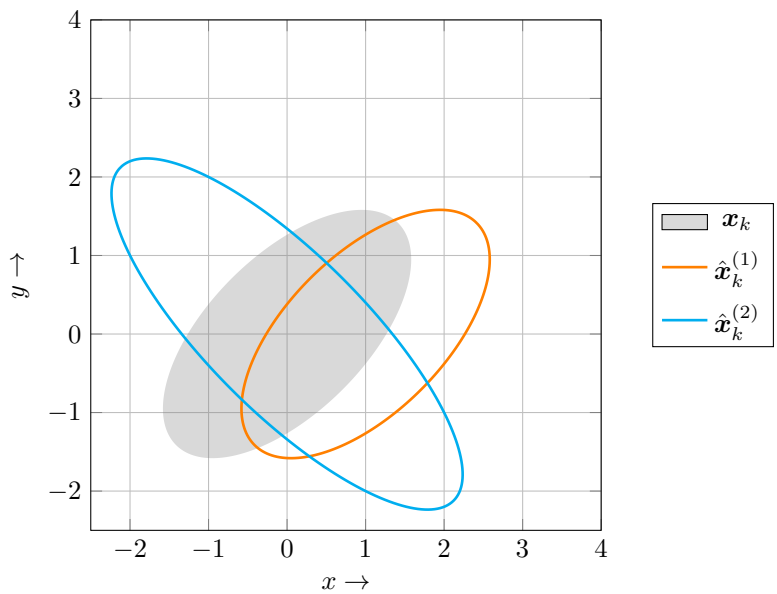


Figure 2.5: This figure gives an example that decoupled measures is used for evaluation. Using the L_p distance between centers, the cyan estimate is better. However, the orange ellipse has a smaller error on the shape estimation.

intersected area divided by the union area, i.e.,

$$\mathrm{IoU}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k) = \frac{\mathrm{Area}(\boldsymbol{x}_k \cap \hat{\boldsymbol{x}}_k)}{\mathrm{Area}(\boldsymbol{x}_k \cup \hat{\boldsymbol{x}}_k)} \quad . \tag{2.7}$$

where $\mathrm{Area}(\cdot)$ is a function that returns the area of a shape.

Even though IoU gives a single score and combines the orientation, size, and center error of two shapes, it has two main limitations in evaluating extended object trackers. Firstly, the intersection and union areas are generally irregular shapes for two elliptical shapes. For elliptical extended object trackers that parameterize elliptical shapes, there is no analytic formula to calculate the area of both intersection and union shapes (see Figure 2.6(a)). It is not critical for computer vision tasks because shapes are typically axes-aligned, and objects are presented using pixels. The intersection and union areas can be easily calculated for axes-aligned rectangles (see Figure 2.6(b)) or tackled by counting pixels in case of irregular areas. Even though the areas can be approximated using sampling, which requires more computation, IoU cannot distinguish two objects when both of them are not overlapping with the ground truth object (see Figure 2.7). This scenario is not typical for computer vision problems but very common for extended object tracking when the objects are manoeuvering.

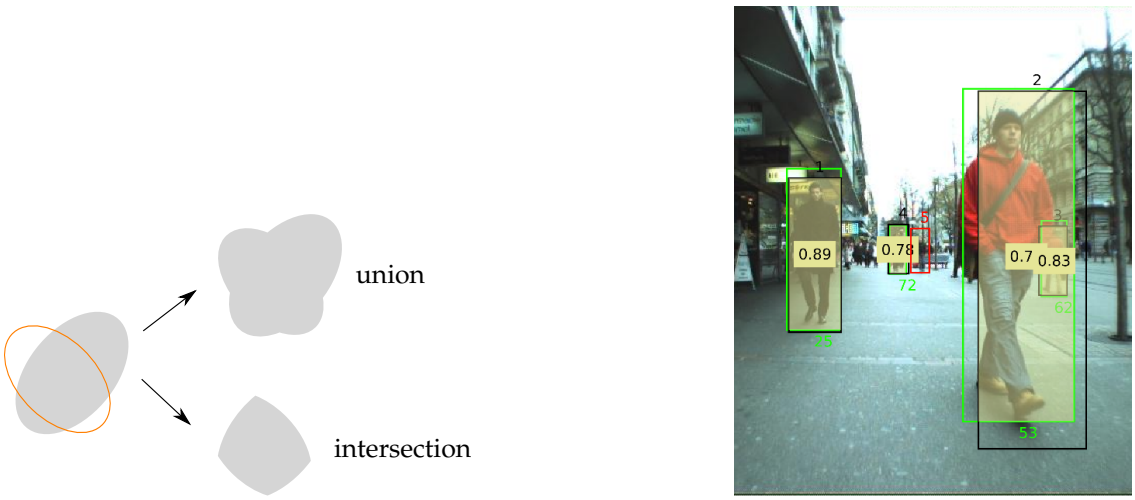


Figure 2.6: (a) shows an example of the irregular shapes of the intersection and union area of two ellipses. (b) gives an example of tracking-by-detection algorithm [53] that uses IoU scores in the evaluation of accuracy. One miss-detected pedestrian is plotted using a red bounding box. The tracked pedestrians and ground truth annotations are green and black bounding boxes, respectively. The overlapping area is patched, and the IoU scores are provided on the image. The green number are the IDs assigned by the tracker, and the ground truth IDs are given using black.

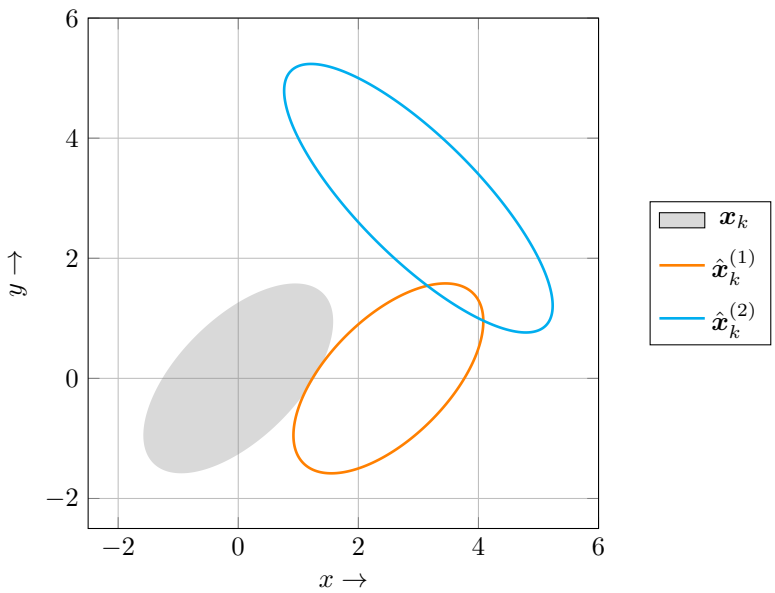


Figure 2.7: A counter-intuitive example of using IoU for extended objects comparison. As neither of the estimates (orange and cyan ellipses) is overlapped with the ground truth (black ellipse), the IoU score of both estimates is 0. However, it is apparent the orange estimate is better as it has less location and shape error.

2.4 Closed-Form Measures for Gaussian Representation

Some geometric shapes can be interpreted as continuous probabilistic distributions. Then, the comparison between shapes can be converted to the distance between distributions. For example, rectangles and ellipses can be described as uniform distributions and Gaussian distributions, respectively.

Representing the shape of an ellipse using a SPD matrix (see (2.5)), an ellipse can be expressed as a Gaussian distribution by representing the center as mean and shape matrix as covariance. Given an extended object state $\boldsymbol{x}_k$ located at $\boldsymbol{m}_k$ with shape matrix $\boldsymbol{\Sigma}_{\boldsymbol{x}_k}$ and an estimate $\hat{\boldsymbol{x}}_k$ located at $\hat{\boldsymbol{m}}_k$ with shape matrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}$, they can be interpreted as Gaussian distributions

$$\mathcal{N}_{\boldsymbol{x}_k} = \mathcal{N}\left(\boldsymbol{m}_k, \boldsymbol{\Sigma}_{\boldsymbol{x}_k}\right), \quad \text{and} \quad \mathcal{N}_{\hat{\boldsymbol{x}}_k} = \mathcal{N}\left(\hat{\boldsymbol{m}}_k, \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}\right) \ . \tag{2.8}$$

Using the Gaussian representation, probabilistic measures can be used to compare ellipses. Among these measures, the Kullback-Leibler divergence, the Hellinger distance, and the Wasserstein distance have closed-form solutions for Gaussian distributions.

2.4.1 Kullback-Leibler Divergence

The Kullback-Leibler divergence was introduced by Solomon Kullback and Richard Leibler in [54] and is also known as relative entropy. Given two continuous distributions $p_1 : \mathbb{R}^d \to \mathbb{R}$ and $p_2 : \mathbb{R}^d \to \mathbb{R}$, Kullback-Leibler divergence is defined as the expectation of the logarithmic difference between these two distributions:

$$d_{\mathrm{KL}}\left(p_1 || p_2\right) = \int p_1(\boldsymbol{\mu}) \log\left(\frac{p_1(\boldsymbol{\mu})}{p_2(\boldsymbol{\mu})}\right) d\boldsymbol{\mu} \ . \tag{2.9}$$

The Kullback-Leibler divergence can be interpreted as the information lost if distribution p_2 is used to approximate p_1. The Kullback-Leibler divergence between two Gaussians has closed-form solution and is defined as

$$d_{\mathrm{KL}}(\mathcal{N}_{\boldsymbol{x}_k} || \mathcal{N}_{\hat{\boldsymbol{x}}_k}) = \frac{1}{2}\left(\mathrm{tr}\left(\boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}^{-1}\boldsymbol{\Sigma}_{\boldsymbol{x}_k}\right) + (\hat{\boldsymbol{m}}_k - \boldsymbol{m}_k)^{\mathrm{T}}\boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}^{-1}(\hat{\boldsymbol{m}}_k - \boldsymbol{m}_k) - d + \ln\left(\frac{\det \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}}{\det \boldsymbol{\Sigma}_{\boldsymbol{x}_k}}\right)\right) \ , \tag{2.10}$$

where d is the dimension of Gaussian distribution, $\mathcal{N}_{\boldsymbol{x}_k}$ and $\mathcal{N}_{\hat{\boldsymbol{x}}_k}$ are defined in (2.8). As two-dimensional shapes are considered in this thesis, we have $d = 2$.

The Kullback-Leibler divergence is non-negative and satisfies the identity property. Nevertheless, it is non-symmetric and does not fulfill triangle inequality. A symmetric measure can be constructed easily as

$$d_{\mathrm{KL}}(\mathcal{N}_{\boldsymbol{x}_k} || \mathcal{N}_{\hat{\boldsymbol{x}}_k}) + d_{\mathrm{KL}}(\mathcal{N}_{\hat{\boldsymbol{x}}_k} || \mathcal{N}_{\boldsymbol{x}_k}) \ . \tag{2.11}$$

The unit of the Kullback-Leibler divergence is natural unit of information, i.e., nats.

2.4.2 Hellinger Distance

The Hellinger distance is a bounded metric that is defined on two distributions. Given two continuous probability distributions $p_1 : \mathbb{R}^d \to \mathbb{R}$ and $p_2 : \mathbb{R}^d \to \mathbb{R}$, the Hellinger distance between them is

$$d_{\text{Hel}}\left(p_1, p_2\right) = \left(\frac{1}{2}\int\left(\sqrt{p_1(\boldsymbol{\mu})} - \sqrt{p_2(\boldsymbol{\mu})}\right)^2 d\boldsymbol{\mu}\right)^{\frac{1}{2}} \quad . \tag{2.12}$$

Expanding (2.12) and applying the fact that the integral of a distribution over its domain is one, we have

$$d_{\text{Hel}}\left(p_1, p_2\right) = \left(1 - \int\sqrt{p_1(\boldsymbol{\mu})p_2(\boldsymbol{\mu})}d\boldsymbol{\mu}\right)^{\frac{1}{2}} \quad . \tag{2.13}$$

According to (2.13), it is obvious that Hellinger distance is bounded from zero to one. Having ground truth object and an estimate expressed in (2.8), the squared Hellinger distance is

$$d^2_{\text{Hel}}\left(\mathcal{N}_{\boldsymbol{x}_k}, \mathcal{N}_{\hat{\boldsymbol{x}}_k}\right) = 1 - \frac{|\boldsymbol{\Sigma}_{\boldsymbol{x}_k}|^{\frac{1}{4}}|\boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}|^{\frac{1}{4}}}{\left|\frac{\boldsymbol{\Sigma}_{\boldsymbol{x}_k}+\boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}}{2}\right|^{\frac{1}{2}}}\exp\left(-\frac{1}{8}(\boldsymbol{m}_k - \hat{\boldsymbol{m}}_k)^{\mathrm{T}}\left(\frac{\boldsymbol{\Sigma}_{\boldsymbol{x}_k} + \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}}{2}\right)^{-1}(\boldsymbol{m}_k - \hat{\boldsymbol{m}}_k)\right) \quad , \tag{2.14}$$

where $|\boldsymbol{\Sigma}_{\boldsymbol{x}_k}|$ is the determinate of $\boldsymbol{\Sigma}_{\boldsymbol{x}_k}$. The Hellinger distance is a true metric and measures distributional divergence. The Hellinger distance has been incorporated with OSPA and used for elliptical extended objects comparison in [55].

2.4.3 Gaussian Wasserstein Distance

The Wasserstein distance arises from the optimal transport problem and is also known as Monge-Kontorovich distance and earth mover's distance in computer vision tasks. The OSPA distance [30] introduced in Section 2.2 is constructed based on Wasserstein distance. The Wasserstein distance is defined on two probability densities or probability measures. Let $d(\cdot,\cdot)$ be a metric on $\mathbb{R}^d$, the p-th Wasserstein distance [56, 57] between probability densities $p_1 : \mathbb{R}^d \to \mathbb{R}$ and $p_2 : \mathbb{R}^d \to \mathbb{R}$ is

$$W_p(p_1, p_2) := \left(\inf_{h\in H(p_1,p_2)}\int d(\boldsymbol{\mu},\boldsymbol{\nu})^p \cdot h(\boldsymbol{\mu},\boldsymbol{\nu})\, d\boldsymbol{\mu}\, d\boldsymbol{\nu}\right)^{1/p} \quad , \tag{2.15}$$

where $H(p_1, p_2)$ denotes the collection of all joint densities $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ that have marginals p_1 and p_2. In general, there are no closed-form solutions for two arbitrary distributions. However, for Gaussian distributions, Wasserstein distance can be calculated analytically. Representing ground truth object and an estimate as (2.8), we have the squared L_2 Gaussian Wasserstein distance as [58]

$$d^2_{\text{GW}}(\mathcal{N}_{\boldsymbol{x}_k}, \mathcal{N}_{\hat{\boldsymbol{x}}_k}) = \| \boldsymbol{m}_k - \hat{\boldsymbol{m}}_k \|^2 + \text{tr}\left(\boldsymbol{\Sigma}_{\boldsymbol{x}} + \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k} - 2\sqrt{\sqrt{\boldsymbol{\Sigma}_{\boldsymbol{x}_k}}\boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}\sqrt{\boldsymbol{\Sigma}_{\boldsymbol{x}_k}}}\right) \quad . \tag{2.16}$$

Intuitively, Wasserstein distance measures the minimum probability mass that needs to be transferred from one probability density in order to match the other. The Wasserstein distance is a

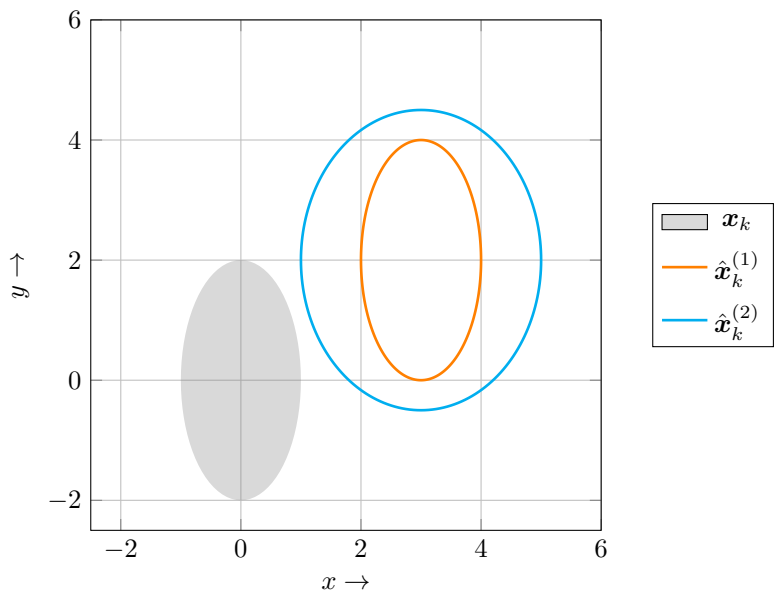


Figure 2.8: A counter-intuitive example of using the Kullback-Leibler divergence and the Hellinger distance to compare ellipses. The filled ellipse depicts ground truth object. Both cyan and orange ellipses have the same center error. The orange ellipse has the same orientation and axes length as the ground truth ellipse. The orange ellipse is expected to have fewer errors.

true metric, i.e., it is symmetric, has the non-negativity, the identity of indiscernible, and the triangle inequality property. By representing elliptical objects as Gaussians, the unit of Gaussian Wasserstein distance in (2.16) is the same as the center, e.g., m, or km.

If two shape matrix $\boldsymbol{\Sigma}_{\boldsymbol{x}_k}$ and $\boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}$ are commutative, i.e., $\boldsymbol{\Sigma}_{\boldsymbol{x}_k}\boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k} = \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}\boldsymbol{\Sigma}_{\boldsymbol{x}_k}$, the squared Gaussian Wasserstein distance boils down to

$$d^2_{\mathrm{GW}}(\mathcal{N}_{\boldsymbol{x}_k}, \mathcal{N}_{\hat{\boldsymbol{x}}_k}) = \| \boldsymbol{m}_k - \hat{\boldsymbol{m}}_k \|^2 + \| \boldsymbol{\Sigma}_{\boldsymbol{x}_k}^{\frac{1}{2}} - \boldsymbol{\Sigma}_{\hat{\boldsymbol{x}}_k}^{\frac{1}{2}} \|_{\mathrm{F}}^2 \ , \tag{2.17}$$

which is called extended square root distance in [59].

2.4.4 Comparison on Closded-Form Probabilistic Measures

In this sub-section, scenarios are constructed to show the arguable comparison results if the Kullback-Leibler divergence and the Hellinger distance are used for evaluating extended objects.

Consider using the Kullback-Leibler divergence, the Hellinger distance, and the Gaussian Wasserstein distance to compare elliptical shapes that are shown in Figure 2.8. Using the parameterization (2.3), the ground truth and two estimates are

$$\boldsymbol{x}_k = \begin{bmatrix} 0, & 0, & 0, & 1, & 2 \end{bmatrix}^{\mathrm{T}}, \quad \hat{\boldsymbol{x}}_k^{(1)} = \begin{bmatrix} 3, & 2, & 0, & 1, & 2 \end{bmatrix}^{\mathrm{T}}, \quad \text{and} \quad \hat{\boldsymbol{x}}_k^{(2)} = \begin{bmatrix} 3, & 2, & 0, & 2, & 2.5 \end{bmatrix}^{\mathrm{T}}.$$

Estimates $\hat{\boldsymbol{x}}_k^{(1)}$ and $\hat{\boldsymbol{x}}_k^{(2)}$ have the same center error but $\hat{\boldsymbol{x}}_k^{(2)}$ has more errors on axes lengths. Therefore, we expect $\hat{\boldsymbol{x}}_k^{(1)}$ has less error than $\hat{\boldsymbol{x}}_k^{(2)}$. The error calculated using the Gaussian Wasserstein distance, the Kullback-Leibler divergence, and the Hellinger distance are

$$d_{\mathrm{GW}}\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(1)}\right) = 5.0990 \ , \qquad d_{\mathrm{KL}}\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(1)}\right) = 5 \ , \qquad d_{\mathrm{Hel}}\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(1)}\right) = 0.8447 \ ,$$
$$d_{\mathrm{GW}}\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(2)}\right) = 5.2202 \ , \qquad d_{\mathrm{KL}}\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(2)}\right) = 1.8063 \ , \qquad d_{\mathrm{Hel}}\left(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k^{(2)}\right) = 0.6993 \ .$$

The Kullback-Leibler divergence and the Hellinger distance evaluate $\hat{\boldsymbol{x}}_k^{(2)}$ having less error than $\hat{\boldsymbol{x}}_k^{(1)}$. This ordering is counter-intuitive because $\hat{\boldsymbol{x}}_k^{(1)}$ only has center error while $\hat{\boldsymbol{x}}_k^{(2)}$ has the same center error but shape error additionally. A symmetric conversion of the Kullback-Leibler divergence is also calculated for both estimates. It gives a similar error ordering to the Kullback-Leibler divergence and the Hellinger distance. For the seek of compactness, the ordering of the converted symmetric Kull-Leibler divergence is not repeated. The reason is that the Kullback-Leibler divergence and the Hellinger distance weigh the difference between centers by the inverse of the covariance matrix, or matrices in the Hellinger distance, according to their definition (2.10) and (2.14). In this example, the center difference between $\hat{\boldsymbol{x}}_k^{(2)}$ and $\boldsymbol{x}_k$ is weighted less.

Furthermore, an artificial scenario is constructed to illustrate the disadvantage of the Kullback-Leibler divergence and the Hellinger distance in Figure 2.9. In Figure 2.9, 100 hypothetical estimates are simulated. All simulated estimates have the same center error. The first estimate has exactly the same shape as the ground truth. The rest of the estimates are constructed by gradually enlarge both axes lengths to the four times of the respective ground truth axes lengths (see Figure 2.9(a)). The Gaussian Wasserstein distance, the Kullback-Leibler divergence, and the Hellinger distance are plotted in Figure 2.9(b), Figure 2.9(c), and Figure 2.9(d) As the center error is fixed, and the shape error increases, the error should increase monotonically by intuition. However, the error curves calculated by the Kullback-Leibler divergence and the Hellinger distance are U-shaped. It means the errors given by these two measures are reducing as the estimated shapes deviate more from the ground truth in the first phase, which is against the intuition.

Note that many other similarity scores and measures, such as Haussdorf-distance [60], are not discussed in this chapter, due to the fact that either they are not suitable for assessing elliptical EOT performance or lack of closed-form solutions.

2.5 Uniform Wasserstein Distance by Sampling

One disadvantage of using Gaussian distributions to represent elliptical shapes is that the Gaussian distributions concentrate more probability mass around the mean. Instead of Gaussian distribution, associating a uniform distribution with an elliptical shape is worth to be investigated.

Compared with a Gaussian distribution, a uniform distribution on the surface of an ellipse is

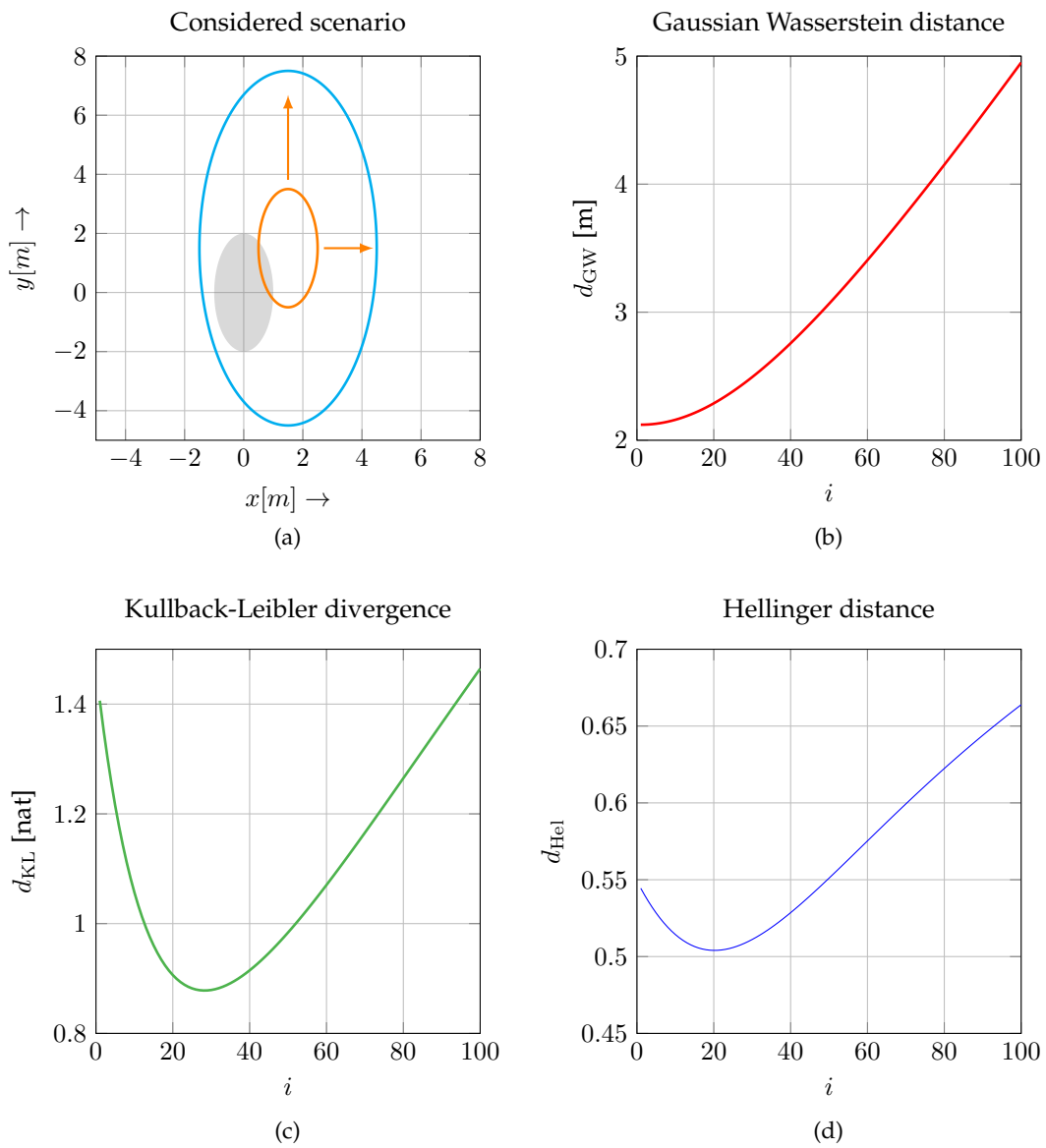


Figure 2.9: The counter-intuitive example of the Kullback-Leibler divergence and the Hellinger distance. The ground truth object $\boldsymbol{x}_k$ is depicted using a filled ellipse. The orange and cyan ellipses in (a) indicate the first estimate $\hat{\boldsymbol{x}}_k^{(1)}$ and the last estimate $\hat{\boldsymbol{x}}_k^{(100)}$, respectively. As the semi-axes lengths of the orange estimate in (a) is becoming larger, the errors evaluated using the Gaussian Wasserstein distance, the Kullback-Leibler divergence, and the Hellinger distance are given in (b), (c), and (d).

a more natural choice for comparing two ellipses. However, there is no analytical solution to solve the Wasserstein distance in this case. Uniformly sampling on the areas of both ground truth object and an estimate can be used as an approximation. Thus, each sample is weighted equally. The OSPA distance between two sets of samples is used to indicate the distance between the ground truth object and the estimate. However, this requires many samples and leads to high computational complexity. Another choice of associating a uniform distribution to an ellipse is a set of uniformly distributed points on the boundary.

Assume n points are selected on the ground truth object $\boldsymbol{x}_k$ and an estimate $\hat{\boldsymbol{x}}_k$, i.e, $\mathbf{P}_{\boldsymbol{x}_k} = \left\{\boldsymbol{\mu}_{\boldsymbol{x}_k}^{(1)}\ ,\ \cdots\ ,\boldsymbol{\mu}_{\boldsymbol{x}_k}^{(n)}\right\}$ and $\mathbf{P}_{\hat{\boldsymbol{x}}_k} = \left\{\boldsymbol{\mu}_{\hat{\boldsymbol{x}}_k}^{(1)}\ ,\ \cdots\ ,\ \boldsymbol{\mu}_{\hat{\boldsymbol{x}}_k}^{(n)}\right\}$. The OSPA distance between these two sets of points is the same as their Wasserstein distance and can be calculated as

$$d_{\mathrm{OSPA}_n}(\mathbf{P}_{\boldsymbol{x}_k}, \mathbf{P}_{\hat{\boldsymbol{x}}_k}) = \min_{\pi \in \Pi} \sqrt{\frac{1}{n}\sum_{i=1}^{n} \parallel \boldsymbol{\mu}_{\boldsymbol{x}_k}^{(i)} - \boldsymbol{\mu}_{\hat{\boldsymbol{x}}}^{(\pi_i)} \parallel^2}\ , \tag{2.18}$$

where Π is the set of permutations of $\{1, \cdots, n\}$. The Gaussian Wasserstein distance and uniform Wasserstein distances might differ in the ordering of two estimates. Figure 2.10 depicts an example of the different orderings of the Gaussian Wasserstein distance and a uniform Wasserstein/OSPA distance. Different numbers of samples on the boundary will result in different distances between the same objects. The different assignments resulted by choosing four and 50 samples on the boundary are shown in Figure 2.11.

Because the density of the samples varies with the size of the object, it is open for discussion of whether the same number of points should be chosen on both the ground truth object and an estimate. In case of the different number of samples are chosen from two objects, the Wasserstein distance does not coincide with a OSPA distance.

Three scenarios are constructed to discuss and compare the Gaussian Wasserstein distance and uniform OSPA distances on the boundary with different numbers of samples. The setting of the constructed scenarios are shown in the first column of Figure 2.12. The corresponding errors evaluated using the Gaussian Wasserstein distance, OSPA distances with four and 100 samples are plotted in the second column of Figure 2.12. Using the parameterization in (2.3), the ground truth object is

$$\mathcal{P}_{\boldsymbol{x}_k} = \begin{bmatrix} 0, & 0, & 0, & 1, & 2 \end{bmatrix}\ . \tag{2.19}$$

For each scenario, 100 estimates with different types of error are simulated.

- *Scenario 1:*
  In the first scenario, the shape variables are assumed to be estimated perfectly. But the error of estimated centers are linearly increasing from 0 to $5\sqrt{2}$, i.e., with

$$\mathcal{P}_{\hat{\boldsymbol{x}}_k^{(1)}} = \begin{bmatrix} 0, & 0, & 0, & 1, & 2 \end{bmatrix}\ ,\ \cdots,\ \mathcal{P}_{\hat{\boldsymbol{x}}_k^{(100)}} = \begin{bmatrix} 5, & 5, & 0, & 1, & 2 \end{bmatrix}\ . \tag{2.20}$$

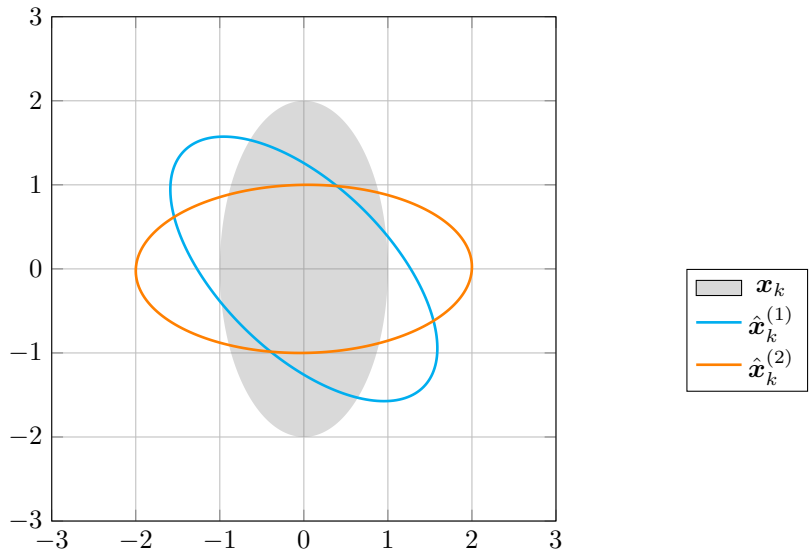


Figure 2.10: Example that shows different ordering for Gaussian Wasserstein distance d_{GW} and a uniform Wasserstein distance d_{OSPA_4}, i.e., four equidistant samples are chosen on the boundary of estimates and ground truth object. The Gaussian Wasserstein distances between the ground truth object $\boldsymbol{x}_k$ and the respective estimate $\hat{\boldsymbol{x}}_k^{(1)}$, $\hat{\boldsymbol{x}}_k^{(2)}$ are 0.98 and 1.4. This means d_{GW} suggests $\hat{\boldsymbol{x}}_k^{(1)}$ has less error than $\hat{\boldsymbol{x}}_k^{(2)}$. The OSPA distances between $\boldsymbol{x}_k$ and $\hat{\boldsymbol{x}}_k^{(1)}$ and $\hat{\boldsymbol{x}}_k^{(2)}$ are 1.2 and 1. Thus, d_{OSPA_4} implies $\hat{\boldsymbol{x}}_k^{(2)}$ is a better estimate than $\hat{\boldsymbol{x}}_k^{(1)}$.

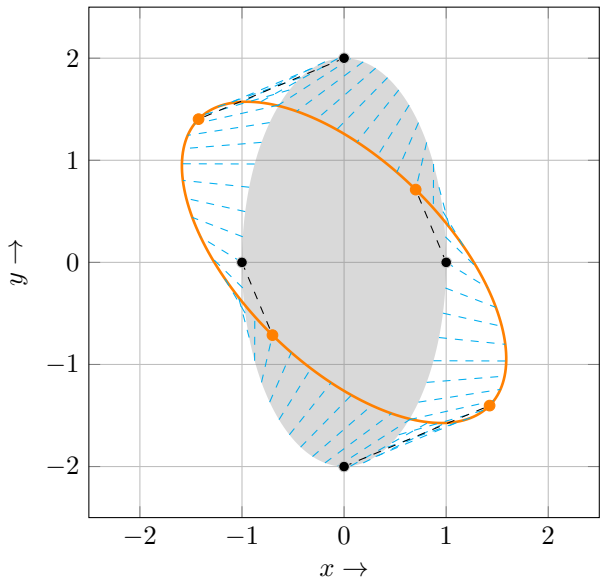


Figure 2.11: Different associations between d_{OPSA_4} and $d_{\mathrm{OSPA}_{50}}$. Cyan dashed lines are the best association between two sets of 50 samples on the boundary, while the black dashed lines indicate the optimal matching between two sets of four points on the boundary of two ellipses. [Yang9] ©2016IEEE

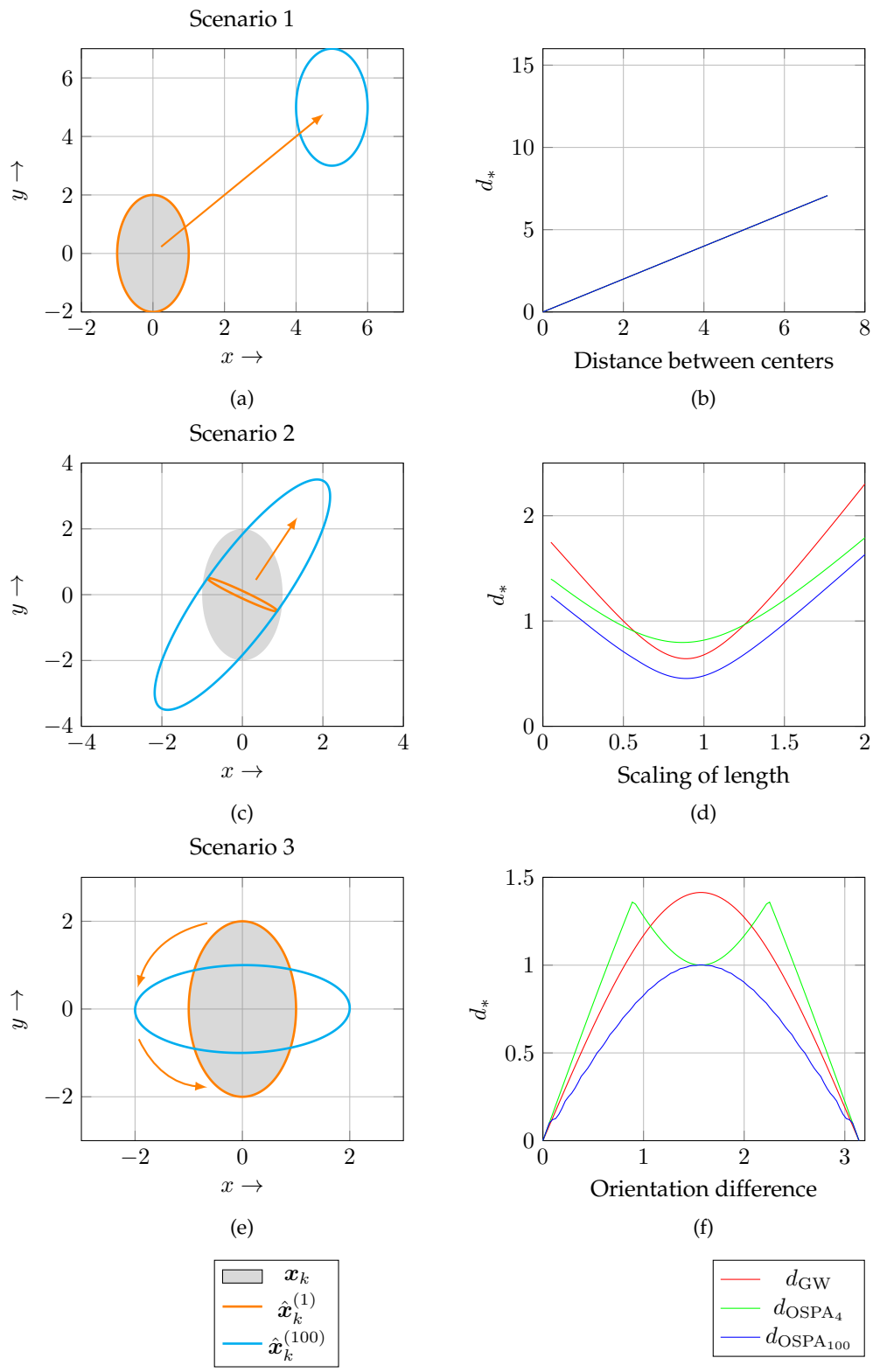


Figure 2.12: The constructed scenarios (left column) and respective Wasserstein distances (right column). The axis label $d_* \in \{d_{\mathrm{GW}}, d_{\mathrm{OSPA}_4}, d_{\mathrm{OSPA}_{100}}\}$.

The setting of the first scenario is plotted in Figure 2.12(a). As it is shown in the error plot in Figure 2.12(b), all Wasserstein distances coincide with each other and grows linearly with the center error. This is consistent with the Euclidean distance that is for evaluating point object trackers.

- *Scenario 2:*
  Given no estimation error on center and a fixed orientation error, the second scenario considers the length of one semi-axes is scaled from 0.1 to 2 times the true length, see Figure 2.12(c). The error plot is given in Figure 2.12(d). Three Wasserstein distances behave similarly, but the Gaussian Wasserstein distance is more sensitive to the errors of lengths compare to the discretized OSPA distances.

- *Scenario 3*:
  In the third scenario, both the location and axes lengths coincide with the ground truth location and lengths, but the orientation error of the estimates range from 0 to π. The setup and error plot are depicted in Figure 2.12(e) and Figure 2.12(f). As the OSPA distance with four samples differs from the Gaussian Wasserstein distance and the discretized OSPA distance with 100 samples, it implies the number of samples that are chosen on the elliptical boundary affects the result of the comparison.

From the constructed scenarios, we can see that the uniformed OSPA distances do not significantly differ from the Gaussian Wasserstein distance. However, it is still open if there exist scenarios such that the Gaussian Wasserstein distance is significantly different from uniform Wasserstein distances. Considering the Gaussian Wasserstein distance is a true metric, has a closed-form solution and intuitive unit, the Gaussian Wasserstein distance is the most suitable metric for comparing elliptical shapes. However, the uniform Wasserstein distance is more general as it can be applied for irregular shapes [61].

Combining the Gaussian Wasserstein distance into the OSPA distance, the metric for multiple elliptical extended object tracking is defined as (2.2) with

$$d^c(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_k) := \min\left(c, d_{\mathrm{GW}}\left(\mathcal{N}_{\boldsymbol{x}_k}, \mathcal{N}_{\hat{\boldsymbol{x}}_k}\right)\right) \quad , \tag{2.21}$$

where c is a user-defined cut-off value, $\mathcal{N}_{\boldsymbol{x}_k}$ and $\mathcal{N}_{\hat{\boldsymbol{x}}_k}$ are the Gaussian representations of the ground truth object and an estimate at time step k, respectively.

2.6 Conclusion

This chapter investigates the metrics that can be used for evaluating elliptical EOT performance. The OSPA distance for evaluating multiple point object trackers is first explained. To adapt OSPA for elliptical extended objects, a distance that can compare elliptical shapes is required. Motivated by this, the possible metrics that can be used for comparing elliptical shapes are overviewed. The

existing measures that are used for comparing elliptical shapes in the literature, such as L_p distance, decoupled measures, and IoU are discussed. After investigating of the shortcomings of using these traditional measures, the closed-form probabilistic measures for Gaussian representation are introduced and compared. By constructing counter-intuitive examples, the Gaussian Wasserstein distance is recommended for assessing elliptical EOT performance. The Gaussian Wasserstein distance is well-known in literature but had not been employed for evaluating extended object trackers before our work [Yang9].

A uniform distribution on the boundary or even the surface is more reasonable and better suited for arbitrarily shaped objects. This chapter further investigates the uniform Wasserstein distances for evaluation by sampling on the elliptical boundary. Three artificial scenarios are constructed, but no significant difference is found between the Gaussian Wasserstein distance and uniform Wasserstein distance approximations for comparing elliptical shapes. The Gaussian Wasserstein distance, in the end, is suggested to be the most suitable measure for comparing elliptical EOT trackers. Since the Gaussian Wasserstein distance was first proposed for elliptical EOT evaluation in [Yang9], it has been used in many EOT literature [33, 34, 35, 36, 37]. Note that the choice of measures for evaluating algorithm performance highly depends on the domain of the tracker and the requirements of the application, hence it is an open topic.

3

Shape Tracking with Explicit Parameterization

Contents

3.1 Related Work 30
3.2 Recursive Bayesian Tracking 31
3.3 Problem Description 34
3.4 Modeling an Elliptical Extended Object 34
3.5 Estimation based on a Tailored Extended Kalman Filter 39
3.6 Evaluation 48
3.7 Discussion 53
3.8 Conclusion 57

This chapter is about a tracker that estimates the kinematics and shape of an extended object given a set of measurements. In this chapter, no clutter is involved, and the set of measurements available is assumed to be generated by the unknown object. Single extended object tracking is a crucial building block of multiple extended objects tracking in many real-world tracking applications.

The tracker presented in this chapter is Multiplicative Error Model-Extended Kalman Filtering* (MEM-EKF*), which provides a closed-form solution for a recursive extended Kalman filter that explicitly estimates the orientation and semi-axes lengths. *MEM* is the abbreviation for the multiplicative error model, which is the key to make the explicit estimation of orientation and axes lengths possible. *EKF* stands for the extended Kalman filter. Due to the high non-linearities, problem tailored approximations are employed in addition to extended Kalman update. Therefore, a "*" is added to denote the specific approximation techniques. As the MEM-EKF* filter directly estimates orientation, axes lengths, and their respective covariance matrices, it is possible to explicitly model the temporal evolution of individual shape parameters and their interdependencies, which is highly relevant to many practical applications. The shape estimation quality of the MEM-EKF* filter is demonstrated using a simulated static ellipse. The strength of modeling the temporal

evolution of shape variables is shown using a constant velocity model and a constant turn model with respect to other state-of-the-art single extended object trackers.

Remark 3.1. This chapter is based on the journal publication [Yang1]. Using a multiplicative error model to relate shape variables and measurement sources for elliptical/rectangular extended objects is first investigated in [Yang10]. The estimation in [Yang10] is achieved by the second-order extended Kalman filter (SOEKF), which requires the calculation of Hessian matrices. An extended Kalman filter tailored to solve this problem is proposed in [Yang7]. The effort of Hessian matrices calculation is saved, but the resulting estimate is sensitive to the prior due to the approximation. In [Yang1], we improved the actual and pseudo-measurement covariance calculation in [Yang7] and included the uncertainty of shape variables in kinematic estimation. The resulting moments approximation in [Yang1] achieved similar performance to second-order extended Kalman filter in [Yang10]. Based on [Yang1], this chapter

- presents a step-wise derivation of the measurement equation,
- adds a more elaborate explanation of the measurement update formulas,
- discusses the connections and differences with random matrix approaches, and
- investigates the pitfalls encountered when selecting parameters.

3.1 Related Work

In this section, a brief discussion on the related work concerning object extent estimation is presented. For more topics on extended object tracking, an overview is given in [4].

Compared to classic point object tracking, the word *extended* in extended object tracking has a two-fold meaning: on the one hand, multiple measurements, which are measurement sources corrupted with sensor noise, are observed for one object; on the other hand, the extent of the object is not negligible and needs to be estimated. Since the EOT originates from classic point object tracking, early studies assume that the measurement sources are located at specific positions of a rigid body shape [15, 62]. Trackers based on this assumption require explicit data association between received measurement and measurement sources. Another model for measurement sources is the spatial distribution model [18, 19]. Without an explicitly specified assignment between measurement source and measurement, the spatial distribution model assumes that each measurement source is an independent random draw from a spatial distribution on the surface, such as a uniform distribution. Furthermore, the number of measurements received at each time step is modeled as a Poisson distribution. The spatial distribution model is widely used in state-of-the-art extended object tracking literature [10, 5, 12, 63].

The object extent can be described with different levels of complexity. Many works in EOT model the extension of the object using geometrical shapes, which are defined by a set of parameters.

The shape of the extended object is modeled as a stick in [43, 19, 64, 65]. There are some other shapes, such as rectangles [12, 66, 67], circles [68, 69, 70], and arbitrary closed contours such as star-convex shapes [11, 71, 61, 27]. One of the seminal works that use an ellipse to approximate the extension is proposed in [72]. The object's extent is defined by a random SPD matrix and is referred to as a random matrix approach. Feldmann *et al.* improved the noise modeling in [5] based on the spatial distribution model of measurement sources. Many works based on [10] and [5] are further developed to adjust to different applications, such as [28] and [46] for extension dynamics, [73, 33] for non-linear measurements, and [74, 75] for a more accurate measurement update. The RHM [11, 71] allows estimating elliptical shapes based on Kalman filters. The shape vector used in RHM is the Cholesky decomposition of the extension matrix. Elliptical object tracking using convolution particle filters is presented in [7]. Other extended object trackers which are based on particle filters, can be found in the overview work [76].

3.2 Recursive Bayesian Tracking

In the framework of Bayesian filtering, the task of object tracking is to estimate the posterior probability of an object state vector given a sequence of noisy detections, i.e., $p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k})$, where $\boldsymbol{x}_k$ denotes the object state and $\boldsymbol{y}_{1:k}$ is the measurements up to time k. Recursive Bayesian tracking consists of two alternating steps: prediction or time update and correction or measurement update. The correction step, which is also known as measurement update, estimates the posterior based on an observed measurement. The measurement at time k is assumed to depend only on the current object state,

$$p(\boldsymbol{y}_k|\boldsymbol{x}_{1:k}) = p(\boldsymbol{y}_k|\boldsymbol{x}_k) \quad . \tag{3.1a}$$

The relationship between measurement and state is reflected in the so-called measurement equation:

$$\boldsymbol{y}_k = f_k(\boldsymbol{x}_x) + \boldsymbol{v}_k \quad , \tag{3.1b}$$

where $\boldsymbol{v}_k$ is additive measurement noise. With the incoming measurement, the posterior of the object state is updated as

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}) = \frac{p\left(\boldsymbol{y}_k|\boldsymbol{x}_k\right) p\left(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1}\right)}{p\left(\boldsymbol{y}_k|\boldsymbol{y}_{1:k-1}\right)} \propto p(\boldsymbol{y}_k|\boldsymbol{x}_k) p\left(\boldsymbol{x}_k|\boldsymbol{y}_{1:k}\right) \quad . \tag{3.1c}$$

The prediction step, which is also referred to as time update, focuses on estimating the state based on the transition or dynamic model. The state probability density is modeled as a first-order Markov chain

$$p(\boldsymbol{x}_k|\boldsymbol{x}_{1:k-1}) = p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) \quad . \tag{3.2a}$$

The transition between two successive states is described by a time-dependent function a_k:

$$\boldsymbol{x}_k = a_k(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_k \ , \tag{3.2b}$$

where $\boldsymbol{w}_k$ is additive process noise. According to the Champman-Kolmogorov equation, the time update formula is

$$p(\boldsymbol{x}_k|\boldsymbol{y}_{1:k-1}) = \int p\left(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}\right) p(\boldsymbol{x}_{k-1}|\boldsymbol{y}_{1:k-1}) d\boldsymbol{x}_{k-1} \ . \tag{3.2c}$$

The prior distribution $p(\boldsymbol{x}_0)$ is assumed to be available.

Solving (3.2c) and (3.1c) is challenging in general as a closed-form solution for the integral and full knowledge of likelihood are needed. However, the Kalman filter gives the minimum mean squared error estimate when the noises and initial states are Gaussian distributed and both motion and measurement models are linear.

3.2.1 Kalman Filter

The Kalman filter maintains the first two moments, i.e., mean and covariance, of the densities and assumes linear measurement and transition models

$$\boldsymbol{y}_k = \mathbf{H}_k \boldsymbol{x}_k + \boldsymbol{v}_k \ , \tag{3.3}$$

$$\boldsymbol{x}_k = \mathbf{A}_k \boldsymbol{x}_{k-1} + \boldsymbol{w}_k \ , \tag{3.4}$$

where $\mathbf{H}_k$ and $\mathbf{A}_k$ are measurement matrix and transition matrix, $\boldsymbol{v}_k$ and $\boldsymbol{w}_k$ are white Gaussian noises with $\boldsymbol{v}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{C}_k^{\boldsymbol{v}}\right)$ and $\boldsymbol{w}_k \sim \mathcal{N}\left(\mathbf{0}, \mathbf{C}_k^{\boldsymbol{w}}\right)$. Given a prior with mean $\hat{\boldsymbol{x}}_0$ and covariance $\mathbf{C}_0^{\boldsymbol{x}}$, the Kalman filter performs measurement and time update recursively. Denoting the updated mean and covariance with a plus sign as superscript, the measurement update is performed as follows:

$$\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k + \mathbf{C}_k^{\boldsymbol{xy}} \left(\mathbf{C}_k^{\boldsymbol{y}}\right)^{-1} \left(\boldsymbol{y}_k - \mathbf{H}\hat{\boldsymbol{x}}_k\right) \ , \tag{3.5}$$

$$\mathbf{C}_k^{\boldsymbol{x}+} = \mathbf{C}_k^{\boldsymbol{x}} - \mathbf{C}_k^{\boldsymbol{xy}} \left(\mathbf{C}_k^{\boldsymbol{y}}\right)^{-1} \left(\mathbf{C}_k^{\boldsymbol{xy}}\right)^{\mathrm{T}} \ , \tag{3.6}$$

where

- $\boldsymbol{y}_k$ is the observed noisy measurement,
- $\hat{\boldsymbol{x}}_k$ is the mean of the predicted state from time $k-1$ to k,
- $\mathbf{C}_k^{\boldsymbol{x}}$ is the predicted covariance of the object's state and indicates the uncertainty of the prediction,

- the covariance for measurement is calculated as

$$\mathbf{C}_k^{\boldsymbol{y}} = \mathbf{H}_k \mathbf{C}_k^{\boldsymbol{x}} \mathbf{H}_k^{\mathrm{T}} + \mathbf{C}_k^{\boldsymbol{v}} \ , \tag{3.7}$$

- the cross covariance between state and measurement is

$$\mathbf{C}_k^{\boldsymbol{xy}} = \mathbf{C}_k^{\boldsymbol{x}} \mathbf{H}_k^{\mathrm{T}} \ . \tag{3.8}$$

Note that $\mathbf{C}_k^{\boldsymbol{xy}} \left(\mathbf{C}_k^{\boldsymbol{y}}\right)^{-1}$ is also referred to as Kalman gain.

The updated mean and covariance are predicted to the next time step as follows:

$$\hat{\boldsymbol{x}}_k = \mathbf{A}_k \hat{\boldsymbol{x}}_{k-1}^{+} \ , \tag{3.9}$$

and

$$\mathbf{C}_k^{\boldsymbol{x}} = \mathbf{A}_k \mathbf{C}_{k-1}^{\boldsymbol{x}+} \mathbf{A}_k^{\mathrm{T}} + \mathbf{C}_k^{\boldsymbol{w}} \ . \tag{3.10}$$

3.2.2 Extended Kalman Filter

Most practical tracking systems deal with non-linear models. The extended Kalman filter is one of the most widely used techniques to solve non-linear estimation problems. The extended Kalman filter linearizes the non-linear measurement or/and dynamic equation(s) using first-order Taylor series expansion and substitutes Jacobian matrices for the corresponding linear matrices in the Kalman filter. Only the non-linear measurement model is considered, as the same contents apply to the non-linear transition model.

Assume the non-linear measurement equation is

$$\boldsymbol{y}_k = f_k(\boldsymbol{x}_k) + \boldsymbol{v}_k \ , \tag{3.11}$$

the Jacobian matrix of f_k is

$$\mathbf{J}_{f_k} = \left.\frac{\partial f_k}{\partial \boldsymbol{x}_k}\right|_{\boldsymbol{x}_k = \hat{\boldsymbol{x}}_k} \ . \tag{3.12}$$

The measurement update follows

$$\hat{\boldsymbol{x}}_k^{+} = \hat{\boldsymbol{x}}_k + \mathbf{C}_k^{\boldsymbol{xy}} \left(\mathbf{C}_k^{\boldsymbol{y}}\right)^{-1} \left(\boldsymbol{y}_k - f_k(\hat{\boldsymbol{x}}_k)\right) \ , \tag{3.13}$$

$$\mathbf{C}_k^{\boldsymbol{x}+} = \mathbf{C}_k^{\boldsymbol{x}} - \mathbf{C}_k^{\boldsymbol{xy}} \left(\mathbf{C}_k^{\boldsymbol{y}}\right)^{-1} \left(\mathbf{C}_k^{\boldsymbol{xy}}\right)^{\mathrm{T}} \ , \tag{3.14}$$

with

$$\mathbf{C}_k^{\boldsymbol{y}} = \mathbf{J}_{f_k} \mathbf{C}_k^{\boldsymbol{x}} \mathbf{J}_{f_k}^{\mathrm{T}} + \mathbf{C}_k^{\boldsymbol{v}}, \tag{3.15}$$

$$\mathbf{C}_k^{\boldsymbol{xy}} = \mathbf{C}_k^{\boldsymbol{x}} \mathbf{J}_{f_k}^{\mathrm{T}} \ . \tag{3.16}$$

3.3 Problem Description

For extended object tracking, we focus on estimating the shape and kinematics of one moving object based on a varying number of measurements. The background clutter is not considered. The state of the object is specified by a state vector $\boldsymbol{x}_k$ with time index k. The paradigm of recursive Bayesian tracking in Section 3.2 is employed except that M_k measurements are observed instead of at most one in the traditional tracking setting. The set of received positional measurements is denoted as $\mathbf{Y}_k = \left\{ \boldsymbol{y}_k^{(i)} \right\}_{i=1}^{M_k}$ with $M_k \geq 0$. The assignment between measurements and object is known, i.e., all the measurements from set $\mathbf{Y}_k$ are generated from the object. Each measurement $\boldsymbol{y}_k^{(i)}$ arises from a measurement source $\boldsymbol{z}_k^{(i)}$ and is perturbed by an additive sensor noise $\boldsymbol{v}_k^{(i)}$. Each measurement source $\boldsymbol{z}_k^{(i)}$ lies on the object surface according to a spatial distribution [18, 19], for example, a uniform distribution.

Estimating the shape of the underlying object based on multiple measurements is significantly more challenging than traditional point object tracking. The difficulties are:

- The measurement of a point object is assumed to be a reflection point of its center, and the main uncertainty is the measurement noise. However, for extended object tracking, besides the measurement noise, each measurement from the extended object can be a reflection of any point on the unknown surface. The measurement sources are distributed differently depending on many factors, such as the types of sensor, the shape of the object, and sensor-object geometry.
- The number of the measurements varies over time and is most often sparse for sensors such as radar.
- The description of object shape requires different levels of complexity. A widely used simplification is to approximate the shape using simple geometric shapes, such as sticks, rectangles, and ellipses. Even for such approximations, the formulation of the shape estimation results in a high dimensional non-linear problem.

3.4 Modeling an Elliptical Extended Object

The Kalman filter requires an explicit measurement equation that relates measurements and the state vector for measurement update. In this section, an elliptical shape parameterization with

orientation and axes lengths is employed. By the introduction of a so-called *multiplicative error*, each observed measurement is related to shape variables, such as orientation and axes lengths.

3.4.1 Parameterization

The kinematic state of the extended object is defined in the same manner as traditional object tracking, i.e.,

$$\boldsymbol{r}_k = \begin{bmatrix} \boldsymbol{m}_k^{\mathrm{T}}, & \dot{\boldsymbol{m}}_k^{\mathrm{T}}, & \dots \end{bmatrix}^{\mathrm{T}} , \tag{3.17}$$

where k is time index, $\boldsymbol{m}_k \in \mathbb{R}^2$ is the location of center, $\dot{\boldsymbol{m}}_k$ is velocity and possible further kinematic quantities, such as turn rate, acceleration. The extent of the object is approximated using an ellipse and parametrized as

$$\boldsymbol{p}_k = \begin{bmatrix} \alpha_k, & l_{k,1}, & l_{k,2} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^3, \tag{3.18}$$

where α_k indicates the orientation, which is defined using the anti-clockwise rotation angle from the x-axis; $l_{k,1}$ and $l_{k,2}$ specify the semi-axes lengths.

Even though parameterization (3.18) is intuitive and widely used in many tracking scenarios [12], the correspondence between elliptical shape and variable representation (3.18) is not one-to-one. To be more specific, one elliptical shape can be represented using multiple shape variable vectors. For example, shape variable vectors $\boldsymbol{p}_k^{(1)} = \begin{bmatrix} \frac{\pi}{3}, & 2, & 1 \end{bmatrix}^{\mathrm{T}}$, $\boldsymbol{p}_k^{(2)} = \begin{bmatrix} \frac{\pi}{3} + \frac{\pi}{2}, & 1, & 2 \end{bmatrix}^{\mathrm{T}}$, and $\boldsymbol{p}_k^{(3)} = \begin{bmatrix} \frac{\pi}{3} + 2\pi, & 2, & 1 \end{bmatrix}^{\mathrm{T}}$ indicate the same elliptical shape. It is important to note that parameterization (3.18) is not limited to elliptical shapes. The rectangular shapes can also use this parameterization, see [Yang10].

3.4.2 Measurement Equation

A measurement equation relates the observed measurement to the unknown object state. In this section, an explicit measurement equation is derived (see Figure 3.1).

First, an axis-aligned ellipse that lies at the origin is considered. The ellipse has semi-axes lengths of $l_{k,1}$ and $l_{k,2}$. For any measurement source $\boldsymbol{z}_k^{(i)}$ that lies on the object surface, there exists $h_{k,1}^{(i)} \in [-1, \quad 1]$ and $h_{k,2}^{(i)} \in [-1, \quad 1]$ such that $\boldsymbol{z}_k^{(i)} = \begin{bmatrix} h_{k,1}^{(i)} l_{k,1}, & h_{k,2}^{(i)} l_{k,2} \end{bmatrix}^{\mathrm{T}}$. Let $\boldsymbol{h}_k^{(i)} = \begin{bmatrix} h_{k,1}^{(i)}, & h_{k,2}^{(i)} \end{bmatrix}^{\mathrm{T}}$, we have

$$\boldsymbol{z}_k^{(i)} = \begin{bmatrix} l_{k,1} & 0 \\ 0 & l_{k,2} \end{bmatrix} \boldsymbol{h}_k^{(i)} . \tag{3.19a}$$

As the measurement sources are assumed to be uniformly distributed on the object extent, the random variable $\boldsymbol{h}_k^{(i)}$ is uniformly distributed on the surface of a unit circle at the origin. Note that if $\boldsymbol{h}_k^{(i)}$ is uniformly distributed in the range $[-1, \quad 1]^2$, a rectangular shape, instead of an elliptical shape, with length $2 \cdot l_{k,1}$ and width $2 \cdot l_{k,2}$, is modeled.

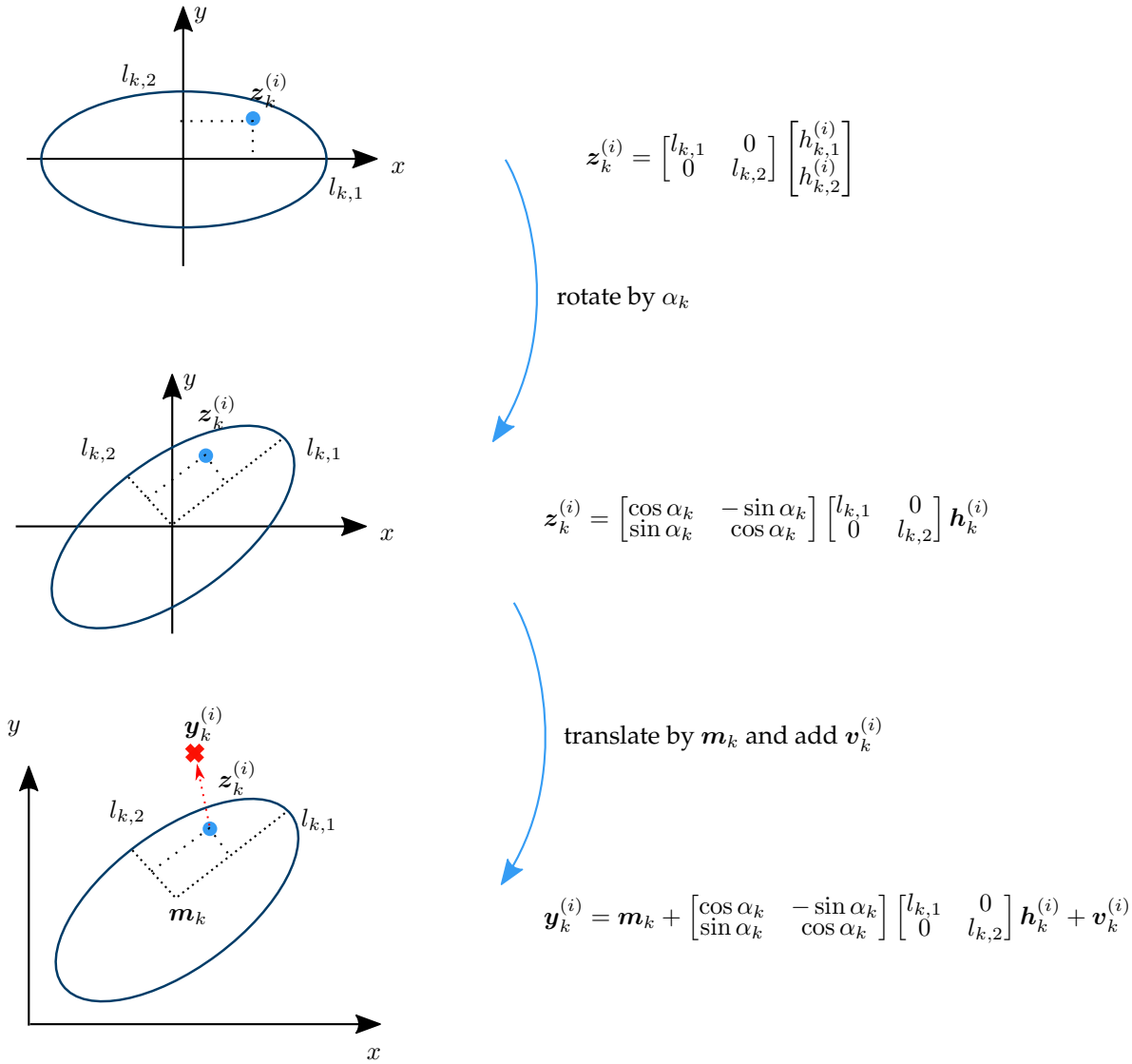


Figure 3.1: This figure shows the derivation of the measurement equation. The key insight is that each measurement $y_k^{(i)}$ is related to shape variable p_k using multiplicative noise $h_k^{(i)}$.

An ellipse with orientation α_k is treated as an axis-aligned ellipse rotated by an angle of α_k. Correspondingly, the measurement source on the extent is also rotated as

$$\boldsymbol{z}_k^{(i)} = \begin{bmatrix} \cos\alpha_k & -\sin\alpha_k \\ \sin\alpha_k & \cos\alpha_k \end{bmatrix} \begin{bmatrix} l_{k,1} & 0 \\ 0 & l_{k,2} \end{bmatrix} \boldsymbol{h}_k^{(i)} \quad . \tag{3.19b}$$

Let matrix $\mathbf{S}_k$ denote $\begin{bmatrix} \cos\alpha_k & -\sin\alpha_k \\ \sin\alpha_k & \cos\alpha_k \end{bmatrix} \begin{bmatrix} l_{k,1} & 0 \\ 0 & l_{k,2} \end{bmatrix}$. Note that matrix $\mathbf{S}_k$ is defined by shape variable $\boldsymbol{p}_k$. Incorporating the coordinates of the center $\boldsymbol{m}_k$ and sensor noise $\boldsymbol{v}_k^{(i)}$, we get

$$\boldsymbol{y}_k^{(i)} = \boldsymbol{m}_k + \mathbf{S}_k \boldsymbol{h}_k^{(i)} + \boldsymbol{v}_k^{(i)} \quad . \tag{3.19c}$$

The formal measurement equation is obtained by relating kinematic state and adding the time index back, which gives us

$$\boldsymbol{y}_k^{(i)} = \mathbf{H}_k \boldsymbol{r}_k + \mathbf{S}_k \boldsymbol{h}_k^{(i)} + \boldsymbol{v}_k^{(i)} \quad , \tag{3.19}$$

where $\mathbf{H}_k$ picks the object location out of the kinematic state.

The random variable $\boldsymbol{h}_k^{(i)}$ is the key to form the explicit measurement equation (3.19). It is treated as a multiplicative error in the estimation process, and this modeling is named Multiplicative Error Model (MEM). The distribution of the multiplicative error is assumed to be Gaussian to facilitate the usage of Kalman filters, i.e., $\boldsymbol{h}_k^{(i)} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}^{\boldsymbol{h}})$ with

$$\mathbf{C}^{\boldsymbol{h}} = \frac{1}{4}\mathbf{I}_2 \tag{3.20}$$

to match the covariance of the circle uniform distribution. Note that $\mathbf{C}^{\boldsymbol{h}}$ plays the similar role as the constant $z = \frac{1}{4}$ in [5].

3.4.3 Dynamic Model

In general, there are no restrictions on the dynamic models for the temporal evolution of the kinematic and shape parameters. As the shape variables of MEM-EKF* are modeled explicitly in a vector, one requires an extra dynamic matrix for shape variables.

Nearly Constant Velocity Model

As the shape estimation in the measurement update is the focus of this chapter, a linear dynamic model is considered. The nearly constant velocity model assumes the dynamics of an extended object is

$$\boldsymbol{r}_{k+1} = \mathbf{A}_{\boldsymbol{r}} \boldsymbol{r}_k + \boldsymbol{w}_k^{\boldsymbol{r}} \quad , \tag{3.21a}$$

$$\boldsymbol{p}_{k+1} = \mathbf{A}_{\boldsymbol{p}} \boldsymbol{p}_k + \boldsymbol{w}_k^{\boldsymbol{p}} \quad , \tag{3.21b}$$

where

- the kinematic state contains location and velocity: $\boldsymbol{r}_k = \begin{bmatrix} \boldsymbol{m}_k, & \dot{\boldsymbol{m}}_k \end{bmatrix}$,
- $\mathbf{A}_{\boldsymbol{r}}$ is the dynamic matrix of kinematic state and is defined as

$$\mathbf{A}_{\boldsymbol{r}} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} , \tag{3.21c}$$

  with the sampling period T ,
- $\mathbf{A}_{\boldsymbol{p}}$ describes the dynamics of shape variable (if the shape of the extended objects is assumed to be fixed over time, $\mathbf{A}_{\boldsymbol{p}}$ is the three dimensional identity matrix and can be omitted),
- $\boldsymbol{w}_k^{\boldsymbol{r}} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{r}}}\right)$ is the process noise of the kinematic state,
- $\boldsymbol{w}_k^{\boldsymbol{p}} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{p}}}\right)$ indicates the process noise of the shape variables.

Coordinated Turn Model

In this parameterization, the kinematic and shape variables are decoupled. There are some motion patterns in which the kinematic state and shape variables are correlated. For example, the object orientation is typically coupled with the turn rate [46] for a maneuvering object. Including the turn rate $\dot{\alpha}_k$ into the kinematic state, the dynamics of the coordinated turn model are

$$\boldsymbol{r}_{k+1} = \mathbf{A}_1\left(\boldsymbol{r}_k\right)\boldsymbol{r}_k + \boldsymbol{w}_k^{\boldsymbol{r}} , \tag{3.22a}$$

$$\boldsymbol{p}_{k+1} = \mathbf{A}_2\boldsymbol{r}_k + \boldsymbol{p}_k + \boldsymbol{w}_k^{\boldsymbol{p}} , \tag{3.22b}$$

where

- $\boldsymbol{r}_k$ defines the kinematic state with turn rate $\dot{\alpha}_k$ appended to (3.17), i.e.,

$$\boldsymbol{r}_k = \begin{bmatrix} \boldsymbol{m}_k, & \dot{\boldsymbol{m}}_k, & \dot{\alpha}_k \end{bmatrix}^{\mathrm{T}} , \tag{3.22c}$$

- the dynamic matrix $\mathbf{A}_1\left(\boldsymbol{r}_k\right)$ takes $\boldsymbol{r}_k$ as parameter:

$$\mathbf{A}_1\left(\boldsymbol{r}_k\right) = \begin{bmatrix} 1 & 0 & \frac{\sin\left(T\dot{\alpha}_k\right)}{\dot{\alpha}_k} & -\frac{1-\cos\left(T\dot{\alpha}_k\right)}{\dot{\alpha}_k} & 0 \\ 0 & 1 & \frac{1-\cos\left(T\dot{\alpha}_k\right)}{\dot{\alpha}_k} & \frac{\sin\left(T\dot{\alpha}_k\right)}{\dot{\alpha}_k} & 0 \\ 0 & 0 & \cos\left(T\dot{\alpha}_k\right) & -\sin\left(T\dot{\alpha}_k\right) & 0 \\ 0 & 0 & \sin\left(T\dot{\alpha}_k\right) & \cos\left(T\dot{\alpha}_k\right) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} , \tag{3.22d}$$

- the matrix

$$\mathbf{A}_2 = \begin{bmatrix} \mathbf{0}_{1\times 4} & T \\ \mathbf{0}_{2\times 4} & \mathbf{0}_{2\times 1} \end{bmatrix} \tag{3.22e}$$

  picks out the turn rate $\dot{\alpha}_k$ from the kinematic state and then adds it to the shape vector in (3.22b).

3.5 Estimation based on a Tailored Extended Kalman Filter

A Kalman filter based on the measurement model (3.19) and dynamic models (3.21a)–(3.22b) is difficult. The main challenges are the following:

a) The estimation is a high-dimensional problem. Compared to point object tracking, the state vector in extended object tracking contains more variables. Apart from the kinematic state, the orientation and two semi-axes lengths need to be estimated.

b) The rotation matrix and introduction of multiplicative noise add more non-linearities to the measurement equation.

c) The mean of the multiplicative noise is zero. If the standard extended Kalman filter is applied, and the Jacobian matrix is evaluated at the estimated value of multiplicative noise (i.e., $\mathbf{0}$), the shape variables are not observable.

d) The kinematic and shape variables are updated separately similar to the random matrix approaches in [5, 28, 46]. By this means, the interdependencies between kinematic and shape variables could not be maintained in the covariance and cross-covariance matrices.

The following solutions are proposed

- Because the shape variables are not observed due to the zero mean multiplicative noise, a pseudo-measurement is used to update shape variables. The deliberately constructed pseudo-measurement facilitates estimation in two aspects. Firstly, the pseudo-measurement is a quadratic conversion of the actual measurement and is uncorrelated to it. Therefore, more information is exploited from the measurement and the shape variables are observable. Secondly, the pseudo-measurement is constructed in a way that its moments can be extracted from the second central moments of the actual measurement.

- To bypass the non-linearities in the measurement equation, linearization is performed on both the actual measurement equation and the pseudo-measurement equation. The extended Kalman filter is performed to calculate the cross-covariance between pseudo-measurement and shape variables. Analytical moment approximations are employed to get the first and second moments of the pseudo-measurement.

- In the process of linearizing the actual measurement equation, the multiplicative noise is treated as as random variable instead of evaluated at its expected value. Therefore, the

spread of the multiplicative noise combined with the uncertainty of shape estimates are kept in the update of the kinematic state.

3.5.1 Sequential Measurement Update

The measurement update in the MEM-EKF* is done sequentially. At each time step k, the predicted prior for kinematic and shape variables are denoted as

$$\hat{\boldsymbol{r}}_k^{(0)},\ \hat{\boldsymbol{p}}_k^{(0)} \quad \text{and} \quad \mathbf{C}_k^{\boldsymbol{r}(0)},\ \mathbf{C}_k^{\boldsymbol{p}(0)} \ .$$

They are updated using each measurement in the set $\{\boldsymbol{y}_k^{(i)}\}_{i=1}^{M_k}$. In the end of the measurement update, the estimates are updated as

$$\hat{\boldsymbol{r}}_k^{(M_k)},\ \hat{\boldsymbol{p}}_k^{(M_k)} \quad \text{and} \quad \mathbf{C}_k^{\boldsymbol{r}(M_k)},\ \mathbf{C}_k^{\boldsymbol{p}(M_k)} \ .$$

For $i = 1, \cdots, M_k$, we denote the estimates that have incorporated the $(i-1)$-th measurement as

$$\hat{\boldsymbol{r}}_k^{(i-1)},\ \hat{\boldsymbol{p}}_k^{(i-1)} \quad \text{and} \quad \mathbf{C}_k^{\boldsymbol{r}(i-1)},\ \mathbf{C}_k^{\boldsymbol{p}(i-1)} \ . \tag{3.23}$$

In the following subsections, only one sequential measurement update is explained, i.e., updating (3.23) using the i-th measurement.

Kinematic State Update

Given the measurement $\boldsymbol{y}_k^{(i)}$, the kinematic variable is updated according to the Kalman filter equations:

$$\bar{\boldsymbol{y}}_k^{(i)} = \mathbf{H}_k \hat{\boldsymbol{r}}_k^{(i-1)} \ , \tag{3.24}$$

$$\hat{\boldsymbol{r}}_k^{(i)} = \hat{\boldsymbol{r}}_k^{(i-1)} + \mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(i)} \left(\mathbf{C}_k^{\boldsymbol{y}(i)}\right)^{-1} \left(\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)}\right) \ , \tag{3.25}$$

$$\mathbf{C}_k^{\boldsymbol{r}(i)} = \mathbf{C}_k^{\boldsymbol{r}(i-1)} - \mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(i)} \left(\mathbf{C}_k^{\boldsymbol{y}(i)}\right)^{-1} \left(\mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(i)}\right)^{\mathrm{T}} \ . \tag{3.26}$$

As the measurement equation is linear in the kinematic state, the cross-correlation $\mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(i)}$ between the measurement and the kinematic state is

$$\mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(i)} = \mathbf{C}_k^{\boldsymbol{r}(i-1)} \mathbf{H}_k^{\mathrm{T}} \ . \tag{3.27}$$

Because of the shape matrix $\mathbf{S}_k$ and multiplicative noise $\boldsymbol{h}_k^{(i)}$, the measurement equation (3.19) is nonlinear in the shape parameters. The challenge is to calculate the covariance of the measurement $\mathbf{C}_k^{\boldsymbol{y}(i)}$.

Algorithm 3.1 Measurement update of the MEM-EKF* filter

Input: Measurements $\{\boldsymbol{y}_k^{(i)}\}_{i=1}^{M_k}$, predicted estimates $\hat{\boldsymbol{r}}_k^{(0)}$, $\hat{\boldsymbol{p}}_k^{(0)}$, $\mathbf{C}_k^{\boldsymbol{r}(0)}$, $\mathbf{C}_k^{\boldsymbol{p}(0)}$, measurement noise covariance $\mathbf{C}_k^{\boldsymbol{v}}$, $\mathbf{H}_k$ as defined in (3.19c), multiplicative noise covariance $\mathbf{C}^{\boldsymbol{h}}$, matrices $\mathbf{F}$ and $\widetilde{\mathbf{F}}$
Output: updated estimates $\hat{\boldsymbol{r}}_k^{(M_k)}$, $\hat{\boldsymbol{p}}_k^{(M_k)}$ and $\mathbf{C}_k^{\boldsymbol{r}(M_k)}$, $\mathbf{C}_k^{\boldsymbol{p}(M_k)}$

For $i = 1, \cdots, M_k$

$$
\begin{aligned}
\begin{bmatrix}\hat{\alpha} & \hat{l}_1 & \hat{l}_2\end{bmatrix}^{\mathrm{T}} &= \hat{\boldsymbol{p}}_k^{(i-1)} \\
\hat{\mathbf{S}} = \begin{bmatrix}\hat{\mathbf{S}}_1 \\ \hat{\mathbf{S}}_2\end{bmatrix} &= \begin{bmatrix}\cos\hat{\alpha} & -\sin\hat{\alpha} \\ \sin\hat{\alpha} & \cos\hat{\alpha}\end{bmatrix}\begin{bmatrix}\hat{l}_1 & 0 \\ 0 & \hat{l}_2\end{bmatrix} \\
\hat{\mathbf{J}}_{\mathbf{S}_1} &= \begin{bmatrix}-\hat{l}_1\sin\hat{\alpha} & \cos\hat{\alpha} & 0 \\ -\hat{l}_2\cos\hat{\alpha} & 0 & -\sin\hat{\alpha}\end{bmatrix} \\
\hat{\mathbf{J}}_{\mathbf{S}_2} &= \begin{bmatrix}\hat{l}_1\cos\hat{\alpha} & \sin\hat{\alpha} & 0 \\ -\hat{l}_2\sin\hat{\alpha} & 0 & \cos\hat{\alpha}\end{bmatrix} \\
\mathbf{C}^{\mathrm{I}} &= \hat{\mathbf{S}}\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{S}}^{\mathrm{T}} \\
\mathbf{C}^{\mathrm{II}} = [\epsilon_{mn}] &= \operatorname{tr}\left(\mathbf{C}_k^{\boldsymbol{p}^{(i-1)}}\hat{\mathbf{J}}_{\mathbf{S}_n}^{\mathrm{T}}\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_m}\right) \text{ for } m, n = 1, 2 \\
\hat{\mathbf{M}} &= \begin{bmatrix} 2\hat{\mathbf{S}}_1\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_1} \\ 2\hat{\mathbf{S}}_2\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_2} \\ \hat{\mathbf{S}}_1\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_2} + \hat{\mathbf{S}}_2\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_1}\end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\bar{\boldsymbol{y}}_k^{(i)} &= \mathbf{H}_k\hat{\boldsymbol{r}}_k^{i-1} \\
\mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(i)} &= \mathbf{C}_k^{\boldsymbol{r}(i-1)}\mathbf{H}_k^{\mathrm{T}} \\
\mathbf{C}_k^{\boldsymbol{y}(i)} &= \mathbf{H}_k\mathbf{C}_k^{\boldsymbol{r}(i-1)}\mathbf{H}_k^{\mathrm{T}} + \mathbf{C}^{\mathrm{I}} + \mathbf{C}^{\mathrm{II}} + \mathbf{C}_k^{\boldsymbol{v}} \\
\boldsymbol{\xi}_k^{(i)} &= \mathbf{F}\left((\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)}) \otimes (\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)})\right) \\
\bar{\boldsymbol{\xi}}_k^{(i)} &= \mathbf{F}\operatorname{vect}\left\{\mathbf{C}_k^{\boldsymbol{y}(i)}\right\} \\
\mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(i)} &= \mathbf{C}_k^{\boldsymbol{p}(i-1)}\hat{\mathbf{M}}^{\mathrm{T}} \\
\mathbf{C}_k^{\boldsymbol{\xi}(i)} &= \mathbf{F}(\mathbf{C}_k^{\boldsymbol{y}(i)} \otimes \mathbf{C}_k^{\boldsymbol{y}(i)})(\mathbf{F} + \widetilde{\mathbf{F}})^{\mathrm{T}}
\end{aligned}
$$

$$
\begin{aligned}
\hat{\boldsymbol{r}}_k^{(i)} &= \hat{\boldsymbol{r}}_k^{(i-1)} + \mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(i)}\left(\mathbf{C}_k^{\boldsymbol{y}(i)}\right)^{-1}\left(\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)}\right) \\
\mathbf{C}_k^{\boldsymbol{r}(i)} &= \mathbf{C}_k^{\boldsymbol{r}(i-1)} - \mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(i)}\left(\mathbf{C}_k^{\boldsymbol{y}(i)}\right)^{-1}\left(\mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(i)}\right)^{\mathrm{T}} \\
\hat{\boldsymbol{p}}_k^{(i)} &= \hat{\boldsymbol{p}}_k^{(i-1)} + \mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(i)}\left(\mathbf{C}_k^{\boldsymbol{\xi}(i)}\right)^{-1}\left(\boldsymbol{\xi}_k^{(i)} - \bar{\boldsymbol{\xi}}_k^{(i)}\right) \\
\mathbf{C}_k^{\boldsymbol{p}(i)} &= \mathbf{C}_k^{\boldsymbol{p}(i-1)} - \mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(i)}\left(\mathbf{C}_k^{\boldsymbol{\xi}(i)}\right)^{-1}\left(\mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(i)}\right)^{\mathrm{T}}
\end{aligned}
$$

End

The time indices k and $(i-1)$ are omitted for the intermediate realizations of variables and matrices, such as $\hat{\alpha}_k^{(i-1)}$, $\hat{l}_{k,1}^{(i-1)}$, $\hat{l}_{k,2}^{(i-1)}$, $\hat{\mathbf{S}}_k^{(i-1)}$, $\hat{\mathbf{J}}_{\mathbf{S}_{k,1}}^{(i-1)}$, $\hat{\mathbf{J}}_{\mathbf{S}_{k,2}}^{(i-1)}$, and $\hat{\mathbf{M}}_k^{(i-1)}$. Source code: `https://github.com/Fusion-Goettingen/`

The first and second row of shape matrix $\mathbf{S}_k$ are

$$\mathbf{S}_{k,1} = \begin{bmatrix} l_{k,1}\cos\alpha_k & -l_2\sin\alpha_k \end{bmatrix} \quad \text{and} \quad \mathbf{S}_{k,2} = \begin{bmatrix} l_{k,1}\sin\alpha_k & l_{k,2}\cos\alpha_k \end{bmatrix} \ . \tag{3.28a}$$

The Jacobian matrices for $\mathbf{S}_{k,1}$ and $\mathbf{S}_{k,2}$ are

$$\mathbf{J}_{\mathbf{S}_{k,1}} = \frac{\partial \mathbf{S}_{k,1}}{\partial \boldsymbol{p}_k} = \begin{bmatrix} -l_{k,1}\sin\alpha_k & \cos\alpha_k & 0 \\ -l_{k,2}\cos\alpha_k & 0 & -\sin\alpha_k \end{bmatrix} \ , \tag{3.28b}$$

$$\mathbf{J}_{\mathbf{S}_{k,2}} = \frac{\partial \mathbf{S}_{k,2}}{\partial \boldsymbol{p}_k} = \begin{bmatrix} l_{k,1}\cos\alpha_k & \sin\alpha_k & 0 \\ -l_{k,2}\sin\alpha_k & 0 & \cos\alpha_k \end{bmatrix} \ , \tag{3.28c}$$

Linearizing $\mathbf{S}_k \boldsymbol{h}_k^{(i)}$ around $\hat{\boldsymbol{p}}_k^{(i-1)}$ and keeping $\boldsymbol{h}_k^{(i)}$ as a random variable, gives us

$$\mathbf{S}_k \boldsymbol{h}_k^{(i)} \approx \underbrace{\hat{\mathbf{S}}_k^{(i-1)} \boldsymbol{h}_k^{(i)}}_{\mathrm{I}} + \underbrace{\begin{bmatrix} \left(\boldsymbol{h}_k^{(i)}\right)^{\mathrm{T}} \hat{\mathbf{J}}_{\mathbf{S}_{k,1}}^{(i-1)} \\ \left(\boldsymbol{h}_k^{(i)}\right)^{\mathrm{T}} \hat{\mathbf{J}}_{\mathbf{S}_{k,2}}^{(i-1)} \end{bmatrix} \left(\boldsymbol{p}_k - \hat{\boldsymbol{p}}_k^{(i-1)}\right)}_{\mathrm{II}} \ , \tag{3.28d}$$

where $\hat{\bullet}^{(i-1)}$ denotes random matrix $\bullet$ realized at the $(i-1)$-th shape estimate $\hat{\boldsymbol{p}}_k^{(i-1)}$. As $\boldsymbol{h}_k^{(i)}$ is zero mean and independent from the shape variables, the first part I and second part II in (3.28d) are uncorrelated. The covariance of $\mathbf{S}_k \boldsymbol{h}_k^{(i)}$ is approximated as the sum of $\mathbf{C}^{\mathrm{I}}$ and $\mathbf{C}^{\mathrm{II}}$, where

$$\mathbf{C}^{\mathrm{I}} = \hat{\mathbf{S}}_k^{(i-1)} \mathbf{C}^{\boldsymbol{h}} (\hat{\mathbf{S}}_k^{(i-1)})^{\mathrm{T}} \ , \tag{3.28e}$$

$$\underbrace{[\epsilon_{mn}]}_{\mathbf{C}^{\mathrm{II}}} = \mathrm{tr}\left(\mathbf{C}_k^{\boldsymbol{p}^{(i-1)}} \left(\hat{\mathbf{J}}_{\mathbf{S}_{k,n}}^{(i-1)}\right)^{\mathrm{T}} \mathbf{C}^{\boldsymbol{h}} \hat{\mathbf{J}}_{\mathbf{S}_{k,m}}^{(i-1)} \right), \quad \text{for } m,n \in \{1,2\} \ . \tag{3.28f}$$

The derivation of (3.28f) is shown in Section 3.5.3. The covariance of the actual measurement is

$$\mathbf{C}_k^{\boldsymbol{y}(i)} = \mathbf{H}_k \mathbf{C}_k^{\boldsymbol{r}(i-1)} \mathbf{H}_k^{\mathrm{T}} + \mathbf{C}^{\mathrm{I}} + \mathbf{C}^{\mathrm{II}} + \mathbf{C}_k^{\boldsymbol{v}} \ . \tag{3.28}$$

Shape Update

As the expected value of the multiplicative error is zero, the shape variables are not observed with a linear estimator [43]. A pseudo-measurement is constructed as follows:

Step 1: Shift the actual measurement by its expected measurement,

$$\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)} \ . \tag{3.29a}$$

Step 2: Take the 2-fold Kronecker product $\otimes$ of the shifted actual measurement,[1]

$$(\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)}) \otimes (\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)}) \ . \tag{3.29b}$$

Step 3: Remove the duplicate element in the resulting Kronecker product

$$\boldsymbol{\xi}_k^{(i)} = \mathbf{F}\left((\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)}) \otimes (\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)})\right) \ , \tag{3.29}$$

where

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \ .$$

The construction of the pseudo-measurement (3.29) is an uncorrelated conversion (c.f. Theorem 3 in [77]). That means the actual measurement and pseudo-measurement are uncorrelated.

Together with the pseudo-measurement $\boldsymbol{\xi}_k^{(i)}$, the shape variables are updated as follows

$$\boldsymbol{p}_k^{(i)} = \hat{\boldsymbol{p}}_k^{(i-1)} + \mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(i)} \left(\mathbf{C}_k^{\boldsymbol{\xi}(i)}\right)^{-1} \left(\boldsymbol{\xi}_k^{(i)} - \bar{\boldsymbol{\xi}}_k^{(i)}\right) \ , \tag{3.30}$$

$$\mathbf{C}_k^{\boldsymbol{p}(i)} = \mathbf{C}_k^{\boldsymbol{p}(i-1)} - \mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(i)} \left(\mathbf{C}_k^{\boldsymbol{\xi}(i)}\right)^{-1} \left(\mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(i)}\right)^{\mathrm{T}} \ , \tag{3.31}$$

where $\bar{\boldsymbol{\xi}}^{(i)}$, $\mathbf{C}_k^{\boldsymbol{\xi}(i)}$, and $\mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(i)}$ are the expected pseudo-measurement, the covariance of the pseudo-measurement, the cross-covariance between the pseudo-measurement and the shape parameters, respectively. The task is to calculate these first and second moments of the pseudo-measurement.

Expected i-th pseudo-measurement:

Suppose the covariance matrix of measurement (3.28) is $\mathbf{C}_k^{\boldsymbol{y}(i)} = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix}$, converting it into a column vector, we have

$$\mathrm{vect}\left\{\mathbf{C}_k^{\boldsymbol{y}(i)}\right\} = \begin{bmatrix} c_{11} & c_{12} & c_{12} & c_{22} \end{bmatrix}^{\mathrm{T}} \ , \tag{3.32a}$$

where the vect-operator converts a matrix into a column vector by stacking its column vectors. By definition, the vectorized covariance matrix (3.32a) is

$$\mathrm{E}\left((\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)}) \otimes (\boldsymbol{y}_k^{(i)} - \bar{\boldsymbol{y}}_k^{(i)})\right) \ , \tag{3.32b}$$

[1]The 2-fold Kronecker product $\otimes$ on a two dimensional vector $\boldsymbol{y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^{\mathrm{T}}$ is

$$\boldsymbol{y} \otimes \boldsymbol{y} = \begin{bmatrix} y_1^2 & y_1 y_2 & y_2 y_1 & y_2^2 \end{bmatrix}^{\mathrm{T}} \ .$$

which leads to the expected value of the pseudo-measurement. Removing the duplicated element gives us the expected i-th pseudo-measurement

$$\bar{\boldsymbol{\xi}}_k^{(i)} = \mathbf{F}\text{vect}\left\{\mathbf{C}_k^{\boldsymbol{y}(i)}\right\} \quad . \tag{3.32}$$

Expected covariance of i-th pseudo-measurement:

The predicted pseudo-measurement covariance is

$$\mathbf{C}_k^{\boldsymbol{\xi}(i)} = \begin{bmatrix} 2c_{11}^2 & 2c_{12}^2 & 2c_{11}c_{12} \\ 2c_{12}^2 & 2c_{22}^2 & 2c_{22}c_{12} \\ 2c_{11}c_{12} & 2c_{22}c_{12} & c_{11}c_{22}+c_{12}^2 \end{bmatrix} \quad , \tag{3.33a}$$

$$= \mathbf{F}(\mathbf{C}_k^{\boldsymbol{y}(i)} \otimes \mathbf{C}_k^{\boldsymbol{y}(i)})(\mathbf{F}+\widetilde{\mathbf{F}})^{\mathrm{T}} \quad . \tag{3.33}$$

where

$$\widetilde{\mathbf{F}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad .$$

The detailed derivation of (3.33a) is shown in Section 3.5.3.

Expected cross-covariance between the i-th pseudo-measurement and shape variables:

The linearization of (3.29) gives us the cross-covariance between the i-th pseudo-measurement and shape variables

$$\mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(i)} = \mathbf{C}_k^{\boldsymbol{p}(i-1)} \left(\hat{\mathbf{M}}_k^{(i-1)}\right)^{\mathrm{T}} \quad , \tag{3.34}$$

with the Jacobian matrix of pseudo-measurement equation (3.29) being

$$\mathbf{M}_k = \begin{bmatrix} 2\mathbf{S}_{k,1}\mathbf{C}^{\boldsymbol{h}}\mathbf{J}_{\mathbf{S}_{k,1}} \\ 2\mathbf{S}_{k,2}\mathbf{C}^{\boldsymbol{h}}\mathbf{J}_{\mathbf{S}_{k,2}} \\ \mathbf{S}_{k,1}\mathbf{C}^{\boldsymbol{h}}\mathbf{J}_{\mathbf{S}_{k,2}} + \mathbf{S}_{k,2}\mathbf{C}^{\boldsymbol{h}}\mathbf{J}_{\mathbf{S}_{k,1}} \end{bmatrix} \quad . \tag{3.34a}$$

The derivation of (3.34a) is shown in Section 3.5.3.

The pseudo-measurement is constructed in a way such that the expected moments of the pseudo-measurement can be built using the centralized second moments of the actual measurements. The pseudo-code of measurement update is given in Algorithm 3.1.

3.5.2 Time Update

With the last step of measurement update, we have estimated $\hat{\boldsymbol{r}}_k^{(M_k)}$, $\hat{\boldsymbol{p}}_k^{(M_k)}$ and their covariance matrices $\mathbf{C}_k^{\boldsymbol{r}(M_k)}$ and $\mathbf{C}_k^{\boldsymbol{p}(M_k)}$, respectively. In the time update step, the updated estimates are predicted into the next time step and indicated with superscript (0) with time index $k+1$.

Constant Velocity Model

As the motion equation is linear in the constant velocity model, the time update is performed in the standard Kalman filter fashion:

$$\hat{\boldsymbol{r}}_{k+1}^{(0)} = \mathbf{A}_{\boldsymbol{r}} \hat{\boldsymbol{r}}_k^{(M_k)} \ , \tag{3.35}$$

$$\mathbf{C}_{k+1}^{\boldsymbol{r}(0)} = \mathbf{A}_{\boldsymbol{r}} \mathbf{C}_k^{\boldsymbol{r}(M_k)} \mathbf{A}_{\boldsymbol{r}}^{\mathrm{T}} + \mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{r}}} \tag{3.36}$$

$$\hat{\boldsymbol{p}}_{k+1}^{(0)} = \mathbf{A}_{\boldsymbol{p}} \hat{\boldsymbol{p}}_k^{(M_k)} \ , \tag{3.37}$$

$$\mathbf{C}_{k+1}^{\boldsymbol{p}(0)} = \mathbf{A}_{\boldsymbol{p}} \mathbf{C}_k^{\boldsymbol{p}(M_k)} \mathbf{A}_{\boldsymbol{p}}^{\mathrm{T}} + \mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{p}}} . \tag{3.38}$$

Coordinated Turn Model

Since the dynamic equation of the kinematic state in the coordinated turn model is non-linear, the kinematic state is predicted using the extended Kalman filter:

$$\hat{\boldsymbol{r}}_{k+1}^{(0)} = \mathbf{A}_1 \left(\hat{\boldsymbol{r}}_k^{(M_k)} \right) \hat{\boldsymbol{r}}_k^{(M_k)} + \hat{\boldsymbol{p}}_k^{(M_k)} \ , \tag{3.39}$$

$$\mathbf{C}_{k+1}^{\boldsymbol{p}(0)} = \mathbf{J}_{\mathbf{A}_1} \mathbf{C}_k^{\boldsymbol{r}(M_k)} \left(\mathbf{J}_{\mathbf{A}_1} \right)^{\mathrm{T}} + \mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{p}}} , \tag{3.40}$$

where the Jacobian matrix of (3.22d) evaluated at the estimate of the last measurement update is

$$\mathbf{J}_{\mathbf{A}_1} = \left. \frac{\partial \mathbf{A}(\boldsymbol{r}_k)}{\partial \boldsymbol{r}_k} \right|_{\boldsymbol{r}_k = \hat{\boldsymbol{r}}_k} . \tag{3.41}$$

The motion equation of shape variables (3.22b) is linear, and updated as

$$\hat{\boldsymbol{p}}_{k+1}^{(0)} = \mathbf{A}_2 \hat{\boldsymbol{r}}_k^{(M_k)} + \hat{\boldsymbol{p}}_k^{(M_k)} \ , \tag{3.42}$$

$$\mathbf{C}_{k+1}^{\boldsymbol{p}(0)} = \mathbf{A}_2 \mathbf{C}_k^{\boldsymbol{r}(M_k)} \mathbf{A}_2^{\mathrm{T}} + \mathbf{C}_k^{\boldsymbol{p}(M_k)} + \mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{p}}} . \tag{3.43}$$

3.5.3 Derivations

This subsection states the derivations of (3.28f), the covariance of pseudo-measurement (3.33a), and the Jacobian matrix of the pseudo-measurement (3.34a). For the sake of compactness, we omit the measurement index (i) and time index k as all the derivations are focused on the i-th sequential measurement update at a time instant k.

The Covariance $\mathbf{C}^{\mathrm{II}}$

Following the definition of covariance, we have

$$\mathbf{C}^{\mathrm{II}} = \mathrm{cov} \left\{ \begin{bmatrix} \boldsymbol{h}^{\mathrm{T}} \hat{\mathbf{J}}_{\mathbf{S}_1} \\ \boldsymbol{h}^{\mathrm{T}} \hat{\mathbf{J}}_{\mathbf{S}_2} \end{bmatrix} (\boldsymbol{p} - \hat{\boldsymbol{p}}) , \begin{bmatrix} \boldsymbol{h}^{\mathrm{T}} \hat{\mathbf{J}}_{\mathbf{S}_1} \\ \boldsymbol{h}^{\mathrm{T}} \hat{\mathbf{J}}_{\mathbf{S}_2} \end{bmatrix} (\boldsymbol{p} - \hat{\boldsymbol{p}}) \right\} = [\epsilon_{mn}] \ . \tag{3.44a}$$

We denote $\tilde{\boldsymbol{p}} = \boldsymbol{p} - \hat{\boldsymbol{p}}$. Both $\tilde{\boldsymbol{p}}$ and $\boldsymbol{h}$ are zero-mean. For each entry ϵ_{mn}, $m, n = 1, 2$, in $\mathbf{C}^{\mathrm{II}}$, we have

$$\epsilon_{mn} = \mathrm{E}\left(\boldsymbol{h}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_m}\tilde{\boldsymbol{p}}\tilde{\boldsymbol{p}}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_n}^{\mathrm{T}}\boldsymbol{h}\right) \quad . \tag{3.44b}$$

As $\boldsymbol{h}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_m}\tilde{\boldsymbol{p}}\tilde{\boldsymbol{p}}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_n}^{\mathrm{T}}\boldsymbol{h}$ is a scalar, we write ϵ_{mn} as

$$\epsilon_{mn} = \mathrm{E}\left(\mathrm{tr}\left(\boldsymbol{h}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_m}\tilde{\boldsymbol{p}}\tilde{\boldsymbol{p}}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_n}^{\mathrm{T}}\boldsymbol{h}\right)\right) \quad . \tag{3.44c}$$

Trace operation is invariant under cyclical permutations, (3.44c) becomes

$$\epsilon_{mn} = \mathrm{E}\left(\mathrm{tr}\left(\tilde{\boldsymbol{p}}\tilde{\boldsymbol{p}}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_n}^{\mathrm{T}}\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_m}\right)\right) \quad . \tag{3.44d}$$

Since trace is a linear operator and can be commuted with expectation, we have

$$\epsilon_{mn} = \mathrm{tr}\left(\mathrm{E}\left(\tilde{\boldsymbol{p}}\tilde{\boldsymbol{p}}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_n}^{\mathrm{T}}\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_m}\right)\right) \quad . \tag{3.44e}$$

The independence between $\boldsymbol{h}$ and $\boldsymbol{p}$ yields

$$\epsilon_{mn} = \mathrm{tr}\left(\mathbf{C}^{\boldsymbol{p}}\hat{\mathbf{J}}_{\mathbf{S}_n}^{\mathrm{T}}\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_m}\right) \quad . \tag{3.44}$$

Covariance of Pseudo-measurement

Given a measurement $\boldsymbol{y} = \begin{bmatrix} y_1, & y_2 \end{bmatrix}^{\mathrm{T}}$ and its expected mean $\bar{\boldsymbol{y}} = \begin{bmatrix} \bar{y}_1, & \bar{y}_2 \end{bmatrix}^{\mathrm{T}}$, the pseudo-measurement is

$$\begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} (y_1 - \bar{y}_1)^2 \\ (y_2 - \bar{y}_2)^2 \\ (y_1 - \bar{y}_1)(y_2 - \bar{y}_2) \end{bmatrix} \quad . \tag{3.45a}$$

The Isserlis's theorem [78] or Wick's theorem [79] calculates higher-order centralized moments of multivariate Gaussian random variables. As the pseudo-measurement is a quadratic form of the actual measurement, the fourth central moments of the actual measurements are required to calculate the covariance of the pseudo-measurement. According to the Isserlis's theorem, we obtain

$$\mathrm{E}\left((\xi_1)^2\right) = 3c_{11}^2 \ , \tag{3.45b}$$

$$\mathrm{E}\left((\xi_2)^2\right) = 3c_{22}^2 \ , \tag{3.45c}$$

$$\mathrm{E}\left(\xi_3\xi_1\right) = 3c_{11}c_{12} \ , \tag{3.45d}$$

$$\mathrm{E}\left(\xi_3\xi_2\right) = 3c_{22}c_{12} \ , \tag{3.45e}$$

$$\mathrm{E}\left((\xi_3)^2\right) = \mathrm{E}\left(\xi_1\xi_2\right) = c_{11}c_{22} + 2c_{12}^2 \ , \tag{3.45f}$$

where c_{mn} denotes $\mathrm{E}\left((y_m - \bar{y}_m)(y_n - \bar{y}_n)\right)$ for $m, n \in \{1, 2\}$.

The mn-th entry of the pseudo-measurement covariance matrix is

$$\operatorname{cov}\{\xi_m, \xi_n\} = \mathrm{E}(\xi_m \xi_n) - \mathrm{E}(\xi_m)\,\mathrm{E}(\xi_n)\,, \text{ for } m, n \in \{1, 2, 3\} \ . \tag{3.45}$$

In the end, the covariance of the pseudo-measurement is obtained as in (3.33a).

Jacobian Matrix of Pseudo-measurement Equation

Denote $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{H}_1$, and $\mathbf{H}_2$ as the first and second row of matrix $\mathbf{S}$ and $\mathbf{H}$, respectively. Expand the pseudo-measurement equation (3.29) as

$$g(\boldsymbol{r}, \boldsymbol{p}) = \begin{bmatrix} (\mathbf{H}_1\boldsymbol{r} + \mathbf{S}_1\boldsymbol{h} + v_1 - \bar{y}_1)^2 \\ (\mathbf{H}_2\boldsymbol{r} + \mathbf{S}_2\boldsymbol{h} + v_2 - \bar{y}_2)^2 \\ (\mathbf{H}_1\boldsymbol{r} + \mathbf{S}_1\boldsymbol{h} + v_1 - \bar{y}_1)(\mathbf{H}_2\boldsymbol{r} + \mathbf{S}_2\boldsymbol{h} + v_2 - \bar{y}_2) \end{bmatrix} \ . \tag{3.46a}$$

Linearizing (3.46a) around $\hat{\boldsymbol{p}}$ and applying the chain rule give us

$$\frac{\partial g}{\partial \boldsymbol{p}} = \begin{bmatrix} 2(\mathbf{H}_1\boldsymbol{r} + \mathbf{S}_1\boldsymbol{h} + v_1 - \bar{y}_1)\,\boldsymbol{h}^{\mathrm{T}}\mathbf{J}_{\mathbf{S}_1} \\ 2(\mathbf{H}_2\boldsymbol{r} + \mathbf{S}_2\boldsymbol{h} + v_2 - \bar{y}_2)\,\boldsymbol{h}^{\mathrm{T}}\mathbf{J}_{\mathbf{S}_2} \\ (\mathbf{H}_1\boldsymbol{r} + \mathbf{S}_1\boldsymbol{h} + v_1 - \bar{y}_1)\,\boldsymbol{h}^{\mathrm{T}}\mathbf{J}_{\mathbf{S}_2} + (\mathbf{H}_2\boldsymbol{r} + \mathbf{S}_2\boldsymbol{h} + v_2 - \bar{y}_2)\,\boldsymbol{h}^{\mathrm{T}}\mathbf{J}_{\mathbf{S}_1} \end{bmatrix} \tag{3.46b}$$

with $\mathbf{J_1}$ and $\mathbf{J_2}$ given in (3.28b) and (3.28c). The cross-covariance between the pseudo-measurement and the shape parameters is approximated as

$$\mathbf{C}^{\boldsymbol{p}\boldsymbol{\xi}} = \operatorname{cov}\left\{ \left.\frac{\partial g}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}=\hat{\boldsymbol{p}}} (\boldsymbol{p} - \hat{\boldsymbol{p}}), \quad \boldsymbol{p} \right\}, \tag{3.46c}$$

$$= \mathbf{C}^{\boldsymbol{p}} \left(\underbrace{\mathrm{E}\left(\left.\frac{\partial g}{\partial \boldsymbol{p}}\right|_{\boldsymbol{p}=\hat{\boldsymbol{p}}} \right)}_{\hat{\mathbf{M}}} \right)^{\mathrm{T}} \ . \tag{3.46d}$$

Evaluating the first row of (3.46b) at $\hat{\boldsymbol{p}}$ gives

$$2(\mathbf{H}_1\boldsymbol{r} - \bar{y}_1)\boldsymbol{h}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_1} + 2\hat{\mathbf{S}}_1\boldsymbol{h}\boldsymbol{h}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_1} + 2v_1\boldsymbol{h}^{\mathrm{T}}\hat{\mathbf{J}}_{\mathbf{S}_1} \ . \tag{3.46e}$$

Taking the expectation of (3.46e) gives us

$$2\hat{\mathbf{S}}_1\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_1} \ . \tag{3.46f}$$

A similar derivation applies to the second and third row of (3.46b). The Jacobian matrix of the

Table 3.1: The ground truth and prior for evaluating moment approximation

	center	shape
Ground Truth	$\boldsymbol{r}_1 = \begin{bmatrix}0, & 0\end{bmatrix}^{\mathrm{T}}$	$\boldsymbol{p}_1 = \begin{bmatrix}\frac{\pi}{3}, & 2, & 9\end{bmatrix}^{\mathrm{T}}$
Prior	$\hat{\boldsymbol{r}}_1^{(0)} = [1, \quad 1]^{\mathrm{T}}$ $\mathbf{C}_1^{\boldsymbol{r}(0)} = \mathrm{diag}\,[1, \quad 1]$	$\hat{\boldsymbol{p}}_1^{(0)} = [0, \quad 2, \quad 12]^{\mathrm{T}}$ $\mathbf{C}_1^{\boldsymbol{p}(0)} = \mathrm{diag}\,[1, \quad 4, \quad 9]$

pseudo-measurement equation is

$$\hat{\mathbf{M}} = \begin{bmatrix} 2\hat{\mathbf{S}}_1\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_1} \\ 2\hat{\mathbf{S}}_2\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_2} \\ \hat{\mathbf{S}}_1\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_2} + \hat{\mathbf{S}}_2\mathbf{C}^{\boldsymbol{h}}\hat{\mathbf{J}}_{\mathbf{S}_1} \end{bmatrix} . \tag{3.46}$$

3.6 Evaluation

In this section, the effectiveness of the extended object tracker MEM-EKF* is demonstrated using simulations. The first simulation considers a static extended object and evaluates the accuracy of the analytic moment approximations proposed in Section 3.5. In the second simulation, a constant velocity model is employed to compare the shape estimation quality of the MEM-EKF* with two other random matrix approaches [5, 44]. The estimation of the turn rate is integrated in the last simulation. The performance of the MEM-EKF* and a random matrix approach [46] that is designed for the coordinated turn model is investigated.

3.6.1 Moment Approximation

In this simulation, the quality of the proposed moment approximations in (3.25), (3.26), (3.30) and (3.31) is compared to the Monte Carlo moment approximation (MEM-MC) and a second-order extended Kalman filter [Yang10] (MEM-SOEKF). Both Monte Carlo and second-order approximations require more computations than MEM-EKF*. Since we want to exclude the effect of the motion model, the extended object is assumed to be static. All three methods use the same prior. Two versions of measurement noise are considered. The ground truth and prior are listed in Table 3.1

The simulation result is shown in Fig. 3.2. As expected, the Monte Carlo moments approximation outperforms the other two methods, whether there is measurement noise or not. In the case of high measurement noise, the MEM-EKF* has a similar approximation quality as the Monte Carlo method. The simulation result in Figure 3.2 shows no significant differences among all these three methods if the number of measurements is sufficient. Thus, one can conclude that the proposed moment approximations have similar quality as Monte Carlo particles even in high noise scenarios.

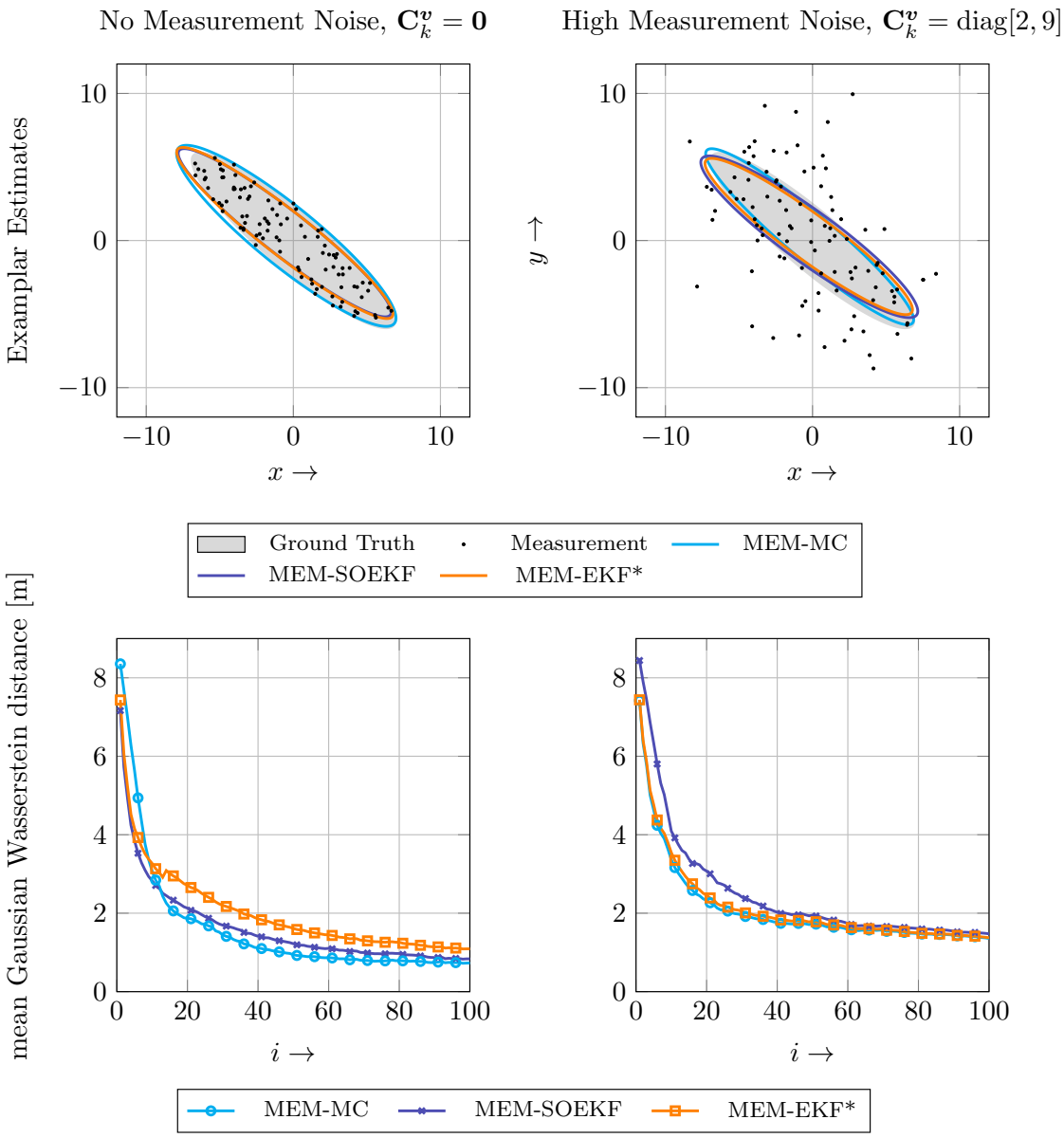


Figure 3.2: Simulation with a static ellipse to evaluate the moments approximation. The left column shows the simulation without measurement noise. The right column describes the simulation with high measurement noise. The first row shows exemplar estimates after 100 measurement updates. The bottom row plots the root mean squared Gaussian Wasserstein distance averaged over 100 runs.

Table 3.2: The prior and parameters for shape variables in the constant velocity simulation. Note that the same notation for the parameters is used as in the original paper.

Parameters	Feldmann *et al.* [5]	Lan *et al.* [28]	MEM-EKF*
Prior of shape	$\begin{bmatrix} 16075 & -13813 \\ -13813 & 32025 \end{bmatrix}$ $\alpha = 50$	$\begin{bmatrix} 16075 & -13813 \\ -13813 & 32025 \end{bmatrix}$ $v = 56$	$\boldsymbol{p}_1^{(0)} = \left[-\frac{\pi}{3}, \quad 200, \quad 90\right]^{\mathrm{T}}$ $\mathbf{C}_1^{\boldsymbol{p}(0)} = \mathrm{diag}\left[1, \quad 490, \quad 490\right]$
Dynamics	$\tau = 50$	$\delta_k = 40$	$\mathbf{C}_{\boldsymbol{p}}^{\boldsymbol{w}} = \mathrm{diag}\left[0.1, \quad 1, \quad 1\right]$

3.6.2 Constant Velocity Model

In this simulation, MEM-EKF* is compared with the random matrix approaches by Feldmann *et al.* [5] and Lan *et al.* [28] to assess the performance of shape estimation.

The size of the tracked object is assumed to be fixed, and only the orientation changes over time. The trajectory is similar to the simulation in [5] and [28] (see Figure 3.3). The extended object has diameters of $340\,\mathrm{m}$ and $80\,\mathrm{m}$. It moves with a velocity of $50\,\mathrm{km/h}$. The number of measurements per time step is Possion distributed with mean 20 as in [28]. The measurement sources are uniformly distributed on the object surface. The covariance of measurement noise is $\mathbf{C}_k^{\boldsymbol{v}} = \mathrm{diag}\left[10000, \quad 400\right]$. As the tracker from Feldmann *et al.* does not model the extension evolution, for a fair comparison, the shape transition matrix is set as $A = \frac{1}{\delta_k}\mathbf{I}_2$ in Lan *et al.* and $\mathbf{A}_k^{\boldsymbol{p}} = \mathbf{I}_3$ in MEM-EKF*.

For a linear motion model, all three trackers estimate the kinematics similarly. The prior of the kinematic state for all tracker is $\left[100, \quad 100, \quad 5, \quad -8\right]^{\mathrm{T}}$ with covariance of $\mathrm{diag}\left[1600, \quad 1600, \quad 16, \quad 16\right]$. The covariance of the process noise is $\mathrm{diag}\left[100, \quad 100, \quad 1, \quad 1\right]$. The prior and parameters for shape/extension are different and are given in Table 3.2.

One example simulation is described in Figure 3.3. As the MEM-EKF* can constrain a low process noise on axes length and allows for comparatively large process noise on orientation, MEM-EKF* outperforms around coordinate turns. The random matrix approach from Feldmann *et al.* detects the changes of the measurement distribution but fails to configure the cause is from orientation or size. The tracker from Lan *et al.* will have better results if more modes of δ_k are incorporated. The estimation error, which combines location and the extent, is plotted as Gaussian Wasserstein distance in Figure 3.4. With single-mode tracker from [28] and using the same linear motion model, both random matrix approaches from Feldmann *et al.* and Lan *et al.* perform similarly.

3.6.3 Coordinated Turn Model

A similar trajectory to [46] is simulated in this section. The true trajectory consists of a constant velocity phase in the first 25 time steps, then a coordinated turn phase in the last 40 time steps (see Figure 3.5). The semi-axes lengths of the simulated object are $170\,\mathrm{m}$ and $40\,\mathrm{m}$. The number

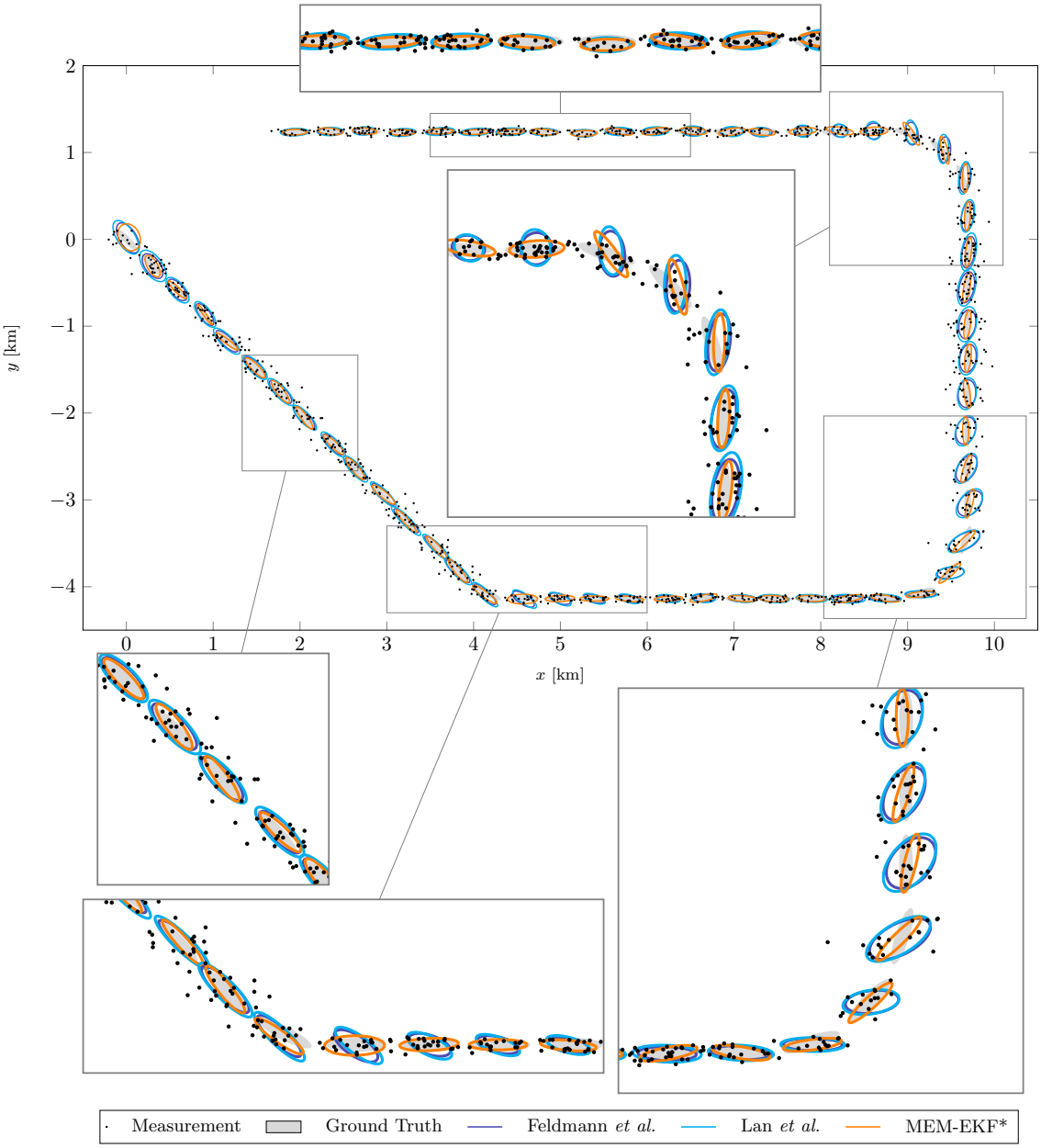


Figure 3.3: The measurements, trajectory, and estimation results of a single example run.

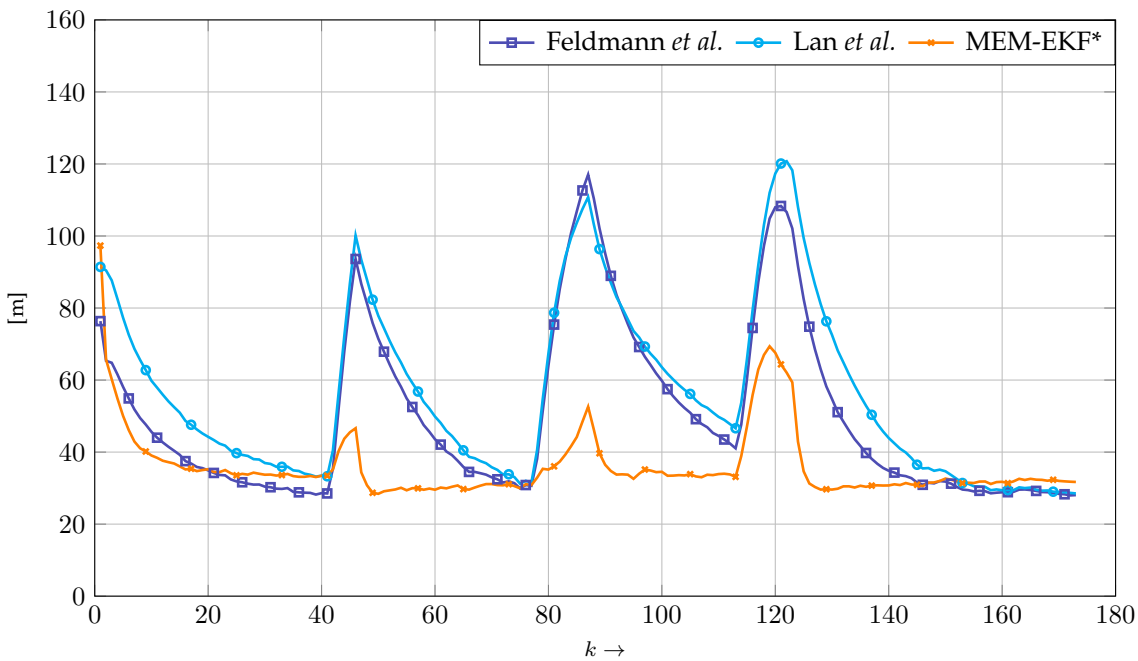


Figure 3.4: Extension and location error based on the root mean squared Gaussian Wasserstein distance. [Yang1] ©2019IEEE

of measurements is Poisson distributed with mean 20 and the measurement noise covariance is $\mathbf{C}_k^{\boldsymbol{v}} = \mathrm{diag}\left[10000, \quad 400\right]$.

Two versions of MEM-EKF* are compared with the M4 approach in [46]. The method M4 improved the prediction step based on the random matrix approach in [5]. It allows for dependencies between the kinematic state and extension matrix. The random matrix approach from [5] is also included as a baseline as it does not incorporate any temporal evolution of the object extent.

A coordinated turn model is used for the object kinematics. The prior for the kinematic state is $\left[100, \quad 100, \quad 100, \quad 20, \quad 0.001\right]^{\mathrm{T}}$ with a covariance of $\mathrm{diag}\left[1600\mathbf{I}_2, \quad 16\mathbf{I}_2, \quad 0.001\right]$. The parameters for the random matrix approaches are $v = 56$, $\tau = 5$, and $T = 1s$. The prior for the MEM-EKF* shape parameter is $\hat{\boldsymbol{p}}_1^{(0)} = \left[\frac{\pi}{3}, \quad 200, \quad 90\right]^{\mathrm{T}}$ with a covariance of $\mathbf{C}_1^{\boldsymbol{p}(0)} = \mathrm{diag}\left[0.2, \quad 360\mathbf{I}_2\right]$. The process noise covariance matrices for the location, velocity, and turn rate are $\mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{r}}} = \mathrm{diag}\left[1000\mathbf{I}_2, \quad 100\mathbf{I}_2, \quad 0.001\right]$.

The evolution of extent can be constrained in the MEM-EKF* using a suitable system noise covariance. Similar to the simulation in Section 3.6.2, the covariance of system noise is tuned in a way such that the axes length are almost fixed over time. To demonstrate the effect of system noise, we

choose two different covariance matrices for process noise:

$$\text{MEM-EKF* (1)} \quad \text{with} \quad \mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{p}}} = \operatorname{diag}\begin{bmatrix}0.01, & \mathbf{I}_2\end{bmatrix} \ , \tag{3.47}$$

$$\text{MEM-EKF* (2)} \quad \text{with} \quad \mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{p}}} = \operatorname{diag}\begin{bmatrix}0.1, & 40\mathbf{I}_2\end{bmatrix} \ . \tag{3.48}$$

Compared to MEM-EKF* (2), MEM-EKF* (1) has low system noise and has more constraint on lengths changes over time. One example estimation result with MEM-EKF*(1), M4 from [46], and Feldmann *et al.* [5] is shown in Figure 3.5.

Figure 3.6 depicts the root mean squared Gaussian Wasserstein distance for this simulation. The performance improves significantly with the turn rate estimation. From Figure 3.6, MEM-EKF* (1) has better performance compared to the other trackers because the axes length changes are restricted similarly as in Section 3.6.2. However, MEM-EKF* (2) gives more estimation error than M4 and MEM-EKF* (1). This means MEM-EKF* is sensitive to the selected parameters and might have worse results (see Figure 3.6).

3.7 Discussion

The random matrix approaches [10, 5] are widely used for tracking elliptical extended objects. Random matrix approaches and MEM-EKF* share many common aspects in modeling the distribution of the measurements, but also have many differences. This section first investigates the relationship between random matrix based elliptical extended object tracking methods and MEM-EKF*. Furthermore, the parameter selection for MEM-EKF* is discussed.

3.7.1 Relationship with Random Matrix Approaches

The measurement likelihoods of the random matrix approach in [5] and the MEM-EKF* are essentially the same. However, these two kinds of shape representation introduce differences in measurement update, time update, and uncertainty representation, etc.

Measurement Likelihood

The random matrix approach in [5] and MEM-EKF* share the same measurement likelihood function. They assume the measurement sources are uniformly distributed on the surface of the extended object, and use a Gaussian distribution to approximate the moments of the elliptical uniform distribution. Denoting the extension matrix as $\boldsymbol{\Sigma}_k$, the measurement likelihood function in [5] is

$$p(\boldsymbol{y}_k^{(i)}|\boldsymbol{r}_k, \boldsymbol{\Sigma}_k) = \mathcal{N}(\boldsymbol{y}_k^{(i)}; \mathbf{H}_k\boldsymbol{r}_k, z\boldsymbol{\Sigma}_k + \mathbf{C}_k^{\boldsymbol{v}}) \ , \tag{3.49}$$

where the constant $z = \frac{1}{4}$ is used to approximate the second order moments of the elliptical uniform distribution. Considering the measurement equation (3.19) and the assumption of Gaussian

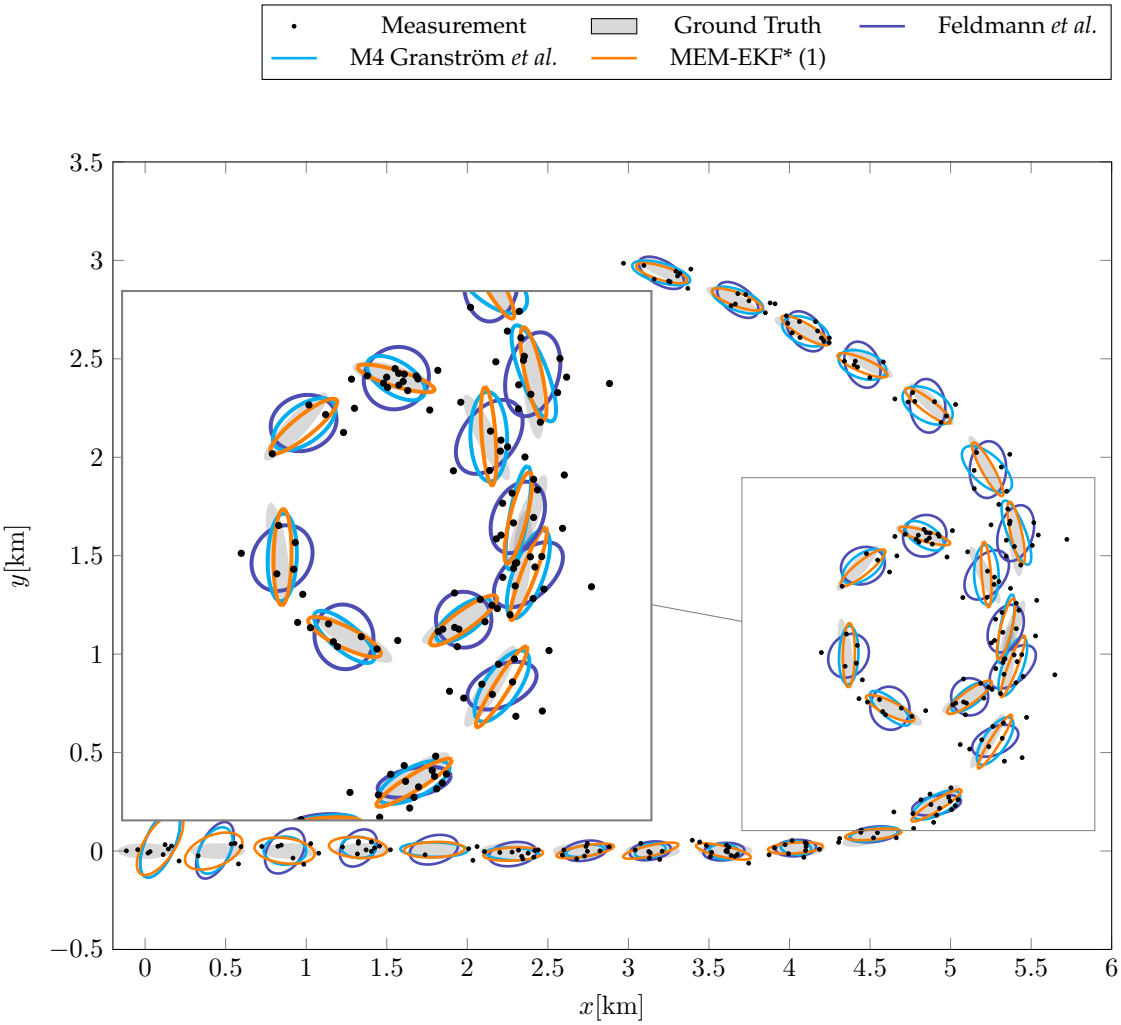


Figure 3.5: The trajectory and estimates of the coordinated turn simulation. For the seek of clear representation, the estimates of MEM-EKF* (2) are omitted as they are worse than M4 from Granström *et al.* and better than Feldmann *et al.* .

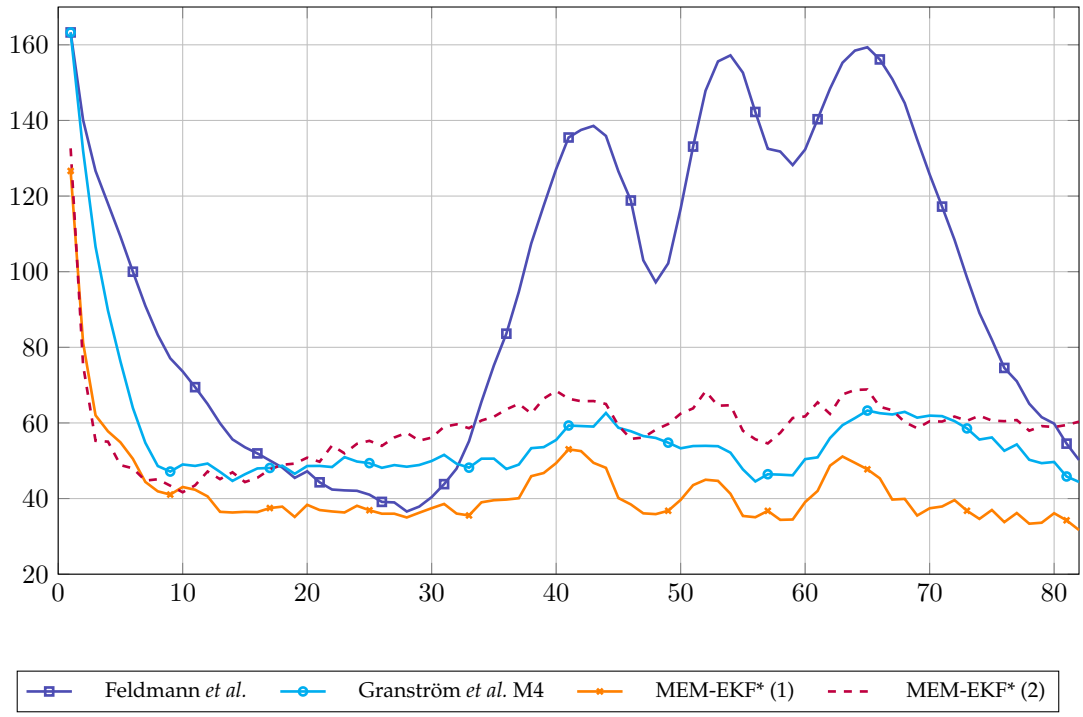


Figure 3.6: Root mean squared Gaussian Wasserstein distance for 100 runs. [Yang1] ©2019IEEE

additive measurement noise together, the measurement likelihood of MEM-EKF* is

$$p(\boldsymbol{y}_k^{(i)}|\boldsymbol{r}_k, \boldsymbol{p}_k) = \mathcal{N}(\boldsymbol{y}_k^{(i)}; \mathbf{H}_k \boldsymbol{r}_k, \mathbf{S}_k \mathbf{C}^{\boldsymbol{h}} \mathbf{S}_k^{\mathrm{T}} + \mathbf{C}_k^{\boldsymbol{v}}) \ , \tag{3.50}$$

where $\mathbf{S}_k \mathbf{S}_k^{\mathrm{T}}$ is extension matrix $\boldsymbol{\Sigma}_k$. Two measurement likelihoods (3.50) and (3.49) are equal as $\mathbf{C}^{\boldsymbol{h}} = \frac{1}{4}\mathbf{I}_2$.

Shape Representation

The physical object extent is represented using a SPD matrix in random matrix approaches while the MEM-EKF* utilizes a three-dimensional shape vector consisting of orientation and semi-axes lengths. These two kinds of shape representations bring the following differences

- Random matrix approaches have a scalar value to indicate the confidence level of the extension estimate. MEM-EKF* maintains a 3×3 covariance matrix. The scalar value used to indicate the confidence level in the random matrix approach is related to (the trace of) the variance of the shape estimate. On one side, the representation of shape estimate uncertainty in MEM-EKF* is more straightforward. The scalar parameter for the uncertainty in random matrix approaches does not allow for a distinction between the uncertainty of size and orientation. On the other side, the confidence indicator in random matrix methods has a

simpler form. A conversion from the covariance matrix to scalar uncertainty introduces a loss of information.

- The time update of MEM-EKF* is more flexible as it can explicitly incorporate the dynamics of shape variables in the motion model. As shown in Section 3.6.2, the fact that the shape changes around coordinate turns are caused by orientation can be expressed by adjusting the process noise of shape variables. This is achieved by adding more orientation process noise and restrict the process noise on axes lengths in the time update. Section 3.6.3 gives a more straightforward example to incorporate the dependency between orientation and turn rate. Depending on the application, other shape dynamics, such as axes lengths changes, can easily be employed into the motion model. The random matrix approach in [5] assumes both orientation and size do not change over time but increase the variance of the extension matrix in the time update. Lan *et al.* expressed the dynamic model of object extension using a transition matrix in [77]. With a multiple mode implementation, the algorithm in [77] allows some degrees of rotation of the extension in the time update. However, the angle of rotation depends on a given parameter not the turn rate. The kinematic state and extension are still assumed to be independent. Granström *et al.* proposed a prediction algorithm, which incorporates the correlation between kinematics and extension [46]. An iterative root-finding algorithm is required in [46].

- The parameters are different (see Table 3.2). For setting the prior of the extension matrix, the random matrix approach in [5] requires an SPD matrix and a scalar, while MEM-EKF* needs a three dimensional mean and a corresponding 3×3 covariance matrix for shape variables. Besides the shape prior, the 3×3 process noise covariance for shape variables needs to be given in MEM-EKF* . Therefore, the MEM-EKF* requires more parameters and is more sensitive to the parameters. More detailed discussion on setting the parameters of MEM-EKF* can be found in Section 3.7.2.

3.7.2 Parameter Selection

Compared to random matrix approaches, the MEM-EKF* requires more parameters for the shape variables (see Table 3.2) and is more sensitive to the selection of parameters.

Two types of shape parameters need to be given for the MEM-EKF*: prior of the first estimate and covariance of the process noise. If an inappropriate set of parameters is chosen, MEM-EKF* might return unreasonable results, such as nearly zero or negative axes lengths. To avoid this, one needs to ensure a comparatively smaller covariance in terms of the mean for shape variables. Either the covariance of the shape variable in the prior or the covariance of the process noise needs to be reasonable. To some extent, the MEM-EKF* is more suitable for applications where most of the object's lengths changes are expressed in the motion model, not in the process noise.

Prior for Shape Variables

For recursive tracking, the prior at the initial time step is usually given and referred to as *first guess*. The prior of the shape variables consists of a three dimensional vector $\hat{\boldsymbol{p}}_1^0$ and a covariance matrix $\mathbf{C}_1^{\boldsymbol{p}(0)}$. As the axes lengths are assumed to be Gaussian distributed, one needs to pay attention that the covariance matrix of the shape variables should not be too large compared to the mean. Otherwise, the lengths might stretch to the negative part of the Gaussian distribution. Note that a high variance of orientation could also raise this problem, as it implies more uncertainties in the shape variables, and these uncertainties will be distributed to lengths of semi-axes in the measurement update. Figure 3.7 depicts examples showing how the estimates are affected by parameters. The ground truth and prior are located at the origin. The shape variables for the ground truth and the mean of prior are

$$\boldsymbol{p} = \begin{bmatrix} \frac{\sqrt{3}}{2}, & 4, & 10 \end{bmatrix}^{\mathrm{T}} \text{ and } \hat{\boldsymbol{p}}_1^{(0)} = \begin{bmatrix} 0.1, & 2, & 12 \end{bmatrix}^{\mathrm{T}}.$$

Five shape covariance matrices for the prior are investigated. All estimates result from the measurement update based on the same 20 measurements and the prior with the same mean but different covariance matrices. The shape covariance $\mathbf{C}_1^{\boldsymbol{p}(0)} = \mathrm{diag}\begin{bmatrix} 0.1, & 4, & 4 \end{bmatrix}$ serves as a benchmark. Its corresponding estimate after 20 times sequential measurement update is indicated by the cyan ellipses in Figure 3.7. The other four estimates and respective prior covariances are described in Figure 3.7(a), 3.7(b), 3.7(c) and 3.7(d). These examples demonstrate that an unreasonably large variance for either orientation or axes lengths gives very small or even negative length estimates.

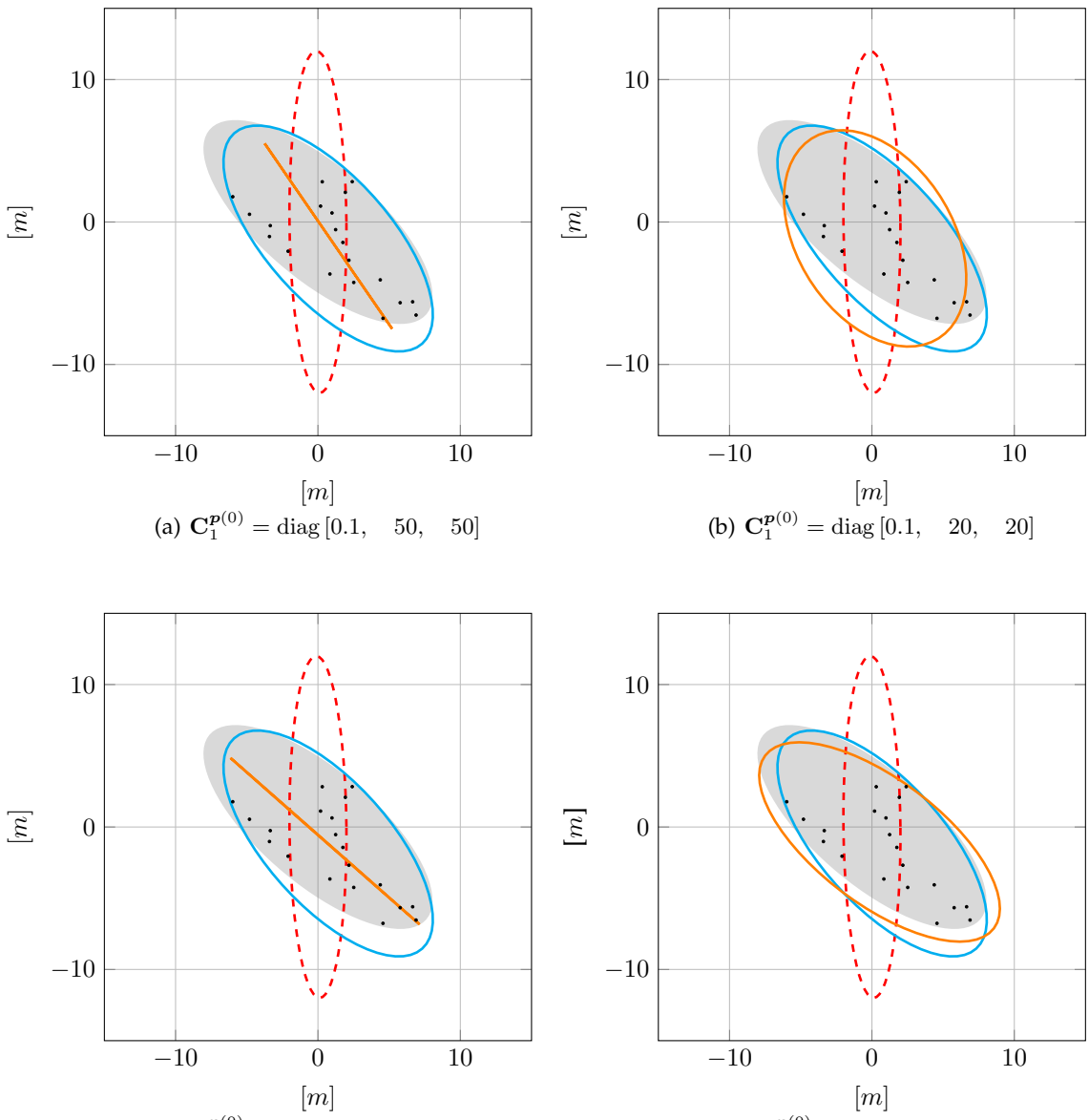
Process noise

As shown in Section 3.6.2, the process noise can be used to restrict the object's size changes and allow more change in the orientation, and vice versa. However, as the covariance of the process noise is added up to the covariance of the shape variables (see (3.38) and (3.43)) and serves as a prior for the next measurement update step, the covariance of the shape process noise needs to assure the lengths are positive in a similar way explained in the previous paragraph.

3.8 Conclusion

In this chapter, an elliptical extended object tracker named MEM-EKF* has been proposed. The main characteristics of the MEM-EKF* are as follows:

- The MEM-EKF* explicitly estimates the orientation and axes lengths of an elliptical object. A 3×3 covariance matrix is employed to indicate the quality.
- The intuitive parameterization using orientation and axes lengths makes it possible to model



(a) $\mathbf{C}_1^{\boldsymbol{p}^{(0)}} = \text{diag}\,[0.1, \quad 50, \quad 50]$

(b) $\mathbf{C}_1^{\boldsymbol{p}^{(0)}} = \text{diag}\,[0.1, \quad 20, \quad 20]$

(c) $\mathbf{C}_1^{\boldsymbol{p}^{(0)}} = \text{diag}\,[0.5, \quad 20, \quad 20]$

(d) $\mathbf{C}_1^{\boldsymbol{p}^{(0)}} = \text{diag}\,[0.5, \quad 4, \quad 4]$

Figure 3.7: The effect of different shape covariance matrices in the prior is depicted in this figure. The ground truth is described as a filled grey ellipse. The mean of the prior is depicted using red dashed ellipses. The cyan estimate serves as a benchmark. The prior covariance matrices used to update the other four orange estimates are indicated in the caption of corresponding sub-figures.

the dynamics of each individual shape variable and their interdependencies. It facilitates many tracking tasks in practical applications, such as maneuvering objects with a fixed size.

- A multiplicative noise is employed to relate each measurement and object state vector. The multiplicative noise introduces more nonlinearities in the measurement equation. The standard linearization of the measurement equation either yields worse results or high complexity. For this reason, problem-tailored techniques such as moment approximation combined with linearization are used to derive the closed-form measurement update in the fashion of the Kalman filter.
- The kinematic state estimate is updated using the actual measurement. However, the shape parameters are updated with a pseudo-measurement constructed from the actual measurement.
- The MEM-EKF* utilizes a sequential measurement update.

The MEM-EKF* deals with single extended object tracking, and focuses on shape estimation using straightforward extension dynamics modeling. A multiple extended object tracking method, which incorporates MEM-EKF* and measurement-to-target data association, is presented in Chapter 4.

4

Multiple Extended Object Tracking without Measurement Partitioning

Contents

4.1 Related Work 63
4.2 Preliminaries 66
4.3 Problem Description 72
4.4 State Prediction 76
4.5 Data Association 77
4.6 State Update 85
4.7 Implementation 86
4.8 Evaluation 92
4.9 Conclusion 103

Chapter 3 presents a single extended object tracker under the assumption that a set of measurements is generated by one object. However, the tracking tasks in many applications deal with multiple extended objects, and the origins of measurements are unknown. One intuitive approach for multiple object tracking is to employ a data association method that decomposes the tracking of multiple objects into a bank of independent single object tracking problems. Hence, the data association method, which aims to figure out the origins of measurements, is a key task in multiple object tracking.

The measurement-to-object assignment in point object tracking is one-to-one. However, the one-to-one constraint is not valid for extended objects as one extended object can generate more than one measurements. Hence, the data association problem for extended object tracking becomes more computational challenging. Most of the state-of-art data association techniques for multiple extended object tracking are decomposed into two steps: i) the measurements partitioning: the

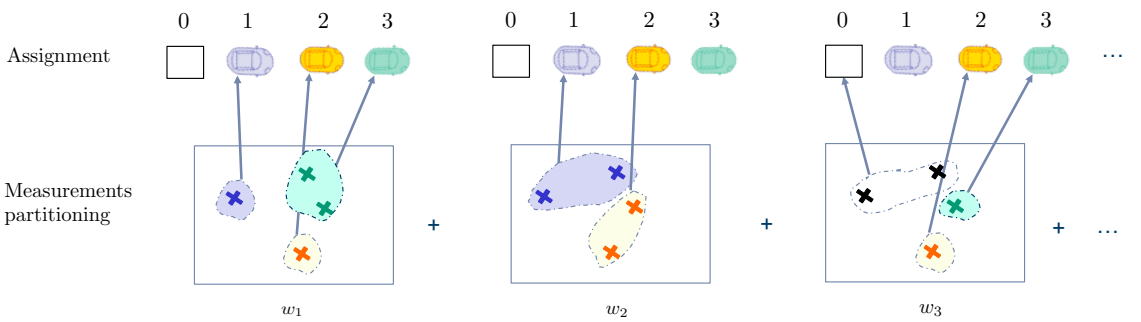


Figure 4.1: An example illustrates the two-step data association method for multiple extended objects. Assume there are four measurements and four measurement sources: background (labeled as 0), object one, object two, and object three. In the first step, the set of measurements is partitioned into measurement cells. Measurements in one measurement cell are indicated using the same color. To reduce the risk of using an erroneous measurement partition, multiple measurement partitions with significant weights are considered. In this example, three measurement partitions are plotted with respective weights w_1, w_2, and w_3. In the second step, the assignment between measurement cells and measurement sources can be achieved using a data association method for point objects.

measurements originating from the same object are grouped into one cell, which is normally achieved by distance-based clustering or the evaluation of all possible measurement-combinations; ii) data association for point objects: associating the measurement cells obtained from the first step to the predicted objects, which is a one-to-one mapping. An example of solving data association for extended objects using two-step approach is shown in Figure 4.1.

An MEOT tracker, which does not require clustering or the enumeration of measurement partitions, is presented in this chapter. The tracker is referred to as *MEM-JIPDA*. As the name suggests, two key components leads to the proposed multiple extended object tracker. First, the MEM-EKF* filter, which is proposed in Chapter 3, is used to estimate the kinematics and extent of each extended object. The zero-inflated PPP [19, 19] is employed to model the measurements. Second, the data association method is inspired by the Joint Integrated Probabilistic Data Association (JIPDA) [80] and adapted for extended objects. The data association result is represented using marginal association probabilities (see Figure 4.2). The exact formulas for marginal association probabilities are achieved by evaluating the likelihoods of all possible assignments.

The efficient data association method is illustrated in an example by assuming Gaussian distributed measurements. Then, the proposed tracker is compared with a state-of-the-art MEOT tracker in a simulation. In the third experiment, the performance of the proposed tracker is illustrated using lidar data to track pedestrians.

Remark 4.1. This chapter is motivated by the journal publication [Yang2], where a multiple extended object tracker that combines the Probability Hypothesis Density (PHD) filter and network flow minimization for labeling and occlusion is proposed. However, the track quality and complexity in [Yang2] highly depend on the clustering method, which is employed in the data association

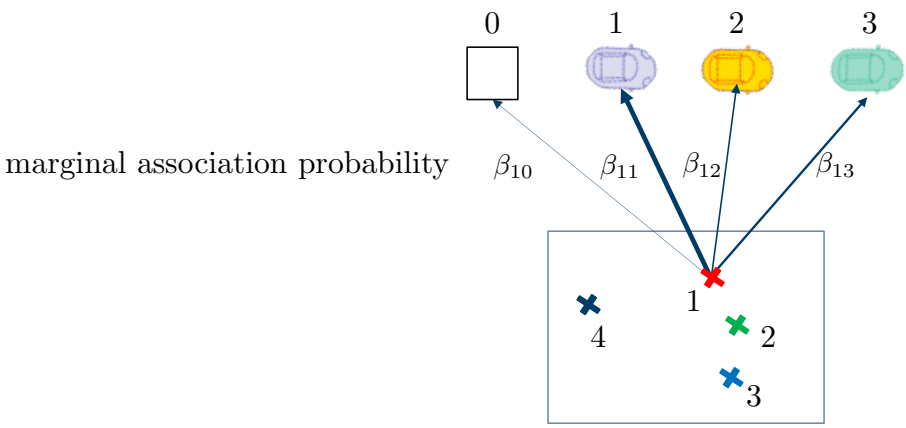


Figure 4.2: An illustrative example of the proposed data association method. The measurement-object assignment is represented using the marginal association probabilities. The β_{10}, β_{11}, β_{12}, β_{13} are the probability of measurement one associated with background, object one, object two, and object three given the existence probability of each measurement source, respectively.

phase. Therefore, a clustering free method is proposed in [Yang5] based on Joint Probabilistic Data Association (JPDA). By assuming all objects are detected, the data association method in [Yang5] is linear in the number of measurements and number of objects. Inspired by JIPDA and combined with the zero-inflated PPP, a likelihood-based data association method for multiple extended object tracking without clustering and enumeration of measurement partitions is proposed in [Yang3]. The data association method in [Yang3] extends [Yang5] by incorporating the detection probability and existence probability for each track. This chapter elaborates an MEOT method, which incorporates the data association method proposed in [Yang3] and the single extended object tracker explained in Chapter 3.

4.1 Related Work

A detailed overview of extended object tracking is given in [4]. The spatial extent estimation for extended object tracking is discussed in Chapter 3. This chapter focuses on multiple object trackers that are decoupled as a set of independent single object trackers by employing data association methods. In this section, we review the state-of-the-art data association methods that are used in multiple point and extended object tracking.

4.1.1 Multiple Point Object Tracking

Data association is one of the most critical and challenging tasks for multiple object trackers. The data association between measurements and objects is often ambiguous because of the presence of clutter, miss-detections, occlusions, and the spatial uncertainty of objects. In general, data association methods can be categorized according to different criteria. A data association method

can be hard or soft. A hard data association method picks the best measurement-object assignment based on an objective function. Then, this assignment is assumed to be correct. A soft data association method associates a measurement to an object with a probability. Correspondingly, the state of an object performs a weighted update using all possible measurements. Data association methods can also be grouped as single-scan and multi-scan methods. As the name suggests, a single-scan method considers the mappings between measurements and objects on the current time step. A multi-scan method associates measurements and objects over a time span. The associations on previous scans can be corrected when a new scan is processed.

Various methods have been developed to solve data association uncertainties. The Nearest Neighbour (NN) and Global Nearest Neighbour (GNN) [81] association methods, which choose the best measurement-object assignment based on a distance measure, are the most intuitive data association approaches. Optimal assignment algorithms, such as the Hungarian method [40], auction algorithm [82], and Murty algorithm [83] are combined with GNN to obtain the optimal association. Instead of choosing only the best association, Probabilistic Data Association (PDA) evaluates the probability of each measurement in the validation region being a correct measurement and performs a weighted update using the corresponding association probabilities. For closely-spaced multiple objects, a measurement might be in the overlapping validation regions of multiple objects. The JPDA filter [84, 31] extends the PDA filter for tracking a fixed number of objects by reasoning all valid joint associations. Both PDA and JPDA assume that the objects exist. The Integrated Probabilistic Data Association (IPDA)[85] and JIPDA [80] are proposed by integrating existence probabilities.

The PDA and JPDA are single-scan data association methods. The Multi-Hypothesis Tracking (MHT) [86] is a batch or multi-scan data association method as it maintains and propagates a subset of association hypotheses over multiple time steps. The MHT method delays the decision for data association and hopes that association ambiguity will be decreased by incorporating data from the future. One of the major disadvantages of MHT is computational complexity. Many MHT variations exist to reduce the computational burden [81, 87]. By relaxing the one-measurement-one-target constraint and assuming the independence across the measurements in the assignment process, a variant of MHT named Probabilistic Multi-Hypothesis Tracking (PMHT) is proposed to reduce the combinatorial complexity. The PMHT converts the combinatorial assignment problem into a continuous optimization problem and makes iterative solutions possible.

Many Random Finite sets (RFS) approaches, which address multiple object tracking, exist. RFS based methods has been very popular in multiple object tracking [88, 17]. RFS theory models objects and measurements as random finite sets, hence allows the generalization of standard Bayesian filter for single object tracking to multiple object tracking. Filters, such as PHD filter [89], Cardinalized Probability Hypothesis Density (CPHD)[90], Labeled Multi-Bernoulli (LMB)[32], and Poisson multi-Bernoulli mixture (PMBM) [35] filter, have been developed.

4.1.2 Multiple Extended Object Tracking

As multiple and unknown number of measurements can be generated by the one extended object, the data association problem for multiple extended object tracking becomes even more challenging compared with point objects as the one-to-one constraint is relaxed. Most multiple extended object trackers bypass many-measurements-to-one-object assignments with *set-to-track* mappings, which are one-to-one. A subset of measurements that is associated with one predicted object as a whole is called a *measurement cell*. The data association is performed between measurement cells and objects. The measurement cells can be obtained by clustering [91, 92, 21], enumeration of possible combinations [93], loopy brief propagation [94], etc. The likelihoods are heuristically incorporated in the clustering process. The approaches that employ clustering techniques are referred to as Clustering and Assignment (C&A) methods. C&A methods are widely used in RFS based multiple extended object trackers, such as PHD [95, 91, 92], LMB[21], and PMBM [35]. The JiFi filter was proposed in [96] in the framework of RFS. It models the measurements generated by each extended object as PPP and assumes the number of tracked objects is known.

A likelihood-based data association method is proposed in [97] and further developed in [98]. Compared with the methods using clustering as a pre-process, likelihood-based data association approaches exploit the spatial measurement model and predicted densities of objects. The likelihood-based data association method in [97] and [98] is solved by stochastic optimization using samples.

Motivated by the many-measurement-to-one-target association problem that is caused by the multi-path propagation in the over-the-horizon radar tracking, Multiple-Detection Probabilistic Data Association (MD-PDA) is presented in [99] and further developed as Multiple-Detection Joint Probabilistic Data Association (MD-JPDA) in [93]. MD-PDA and MD-JPDA assume that there is a maximum number of generated measurements for each object. As there is a maximum of four possible propagation paths for over-the-horizon radars, this number usually is four. MD-JPDA evaluates the Kalman gain by collectively exhausting all possible joint association events in the validation area. Based on IPDA, the many-to-one association is adopted and the Generalized Probabilistic Data Association (GPDA) is proposed in [100]. The Generalized Joint Probabilistic Data Association (GJPDA) is an extension of the GPDA, which adapts the PPP model to interpret the number of measurements generated by extended objects [101]. Both MD-JPDA and GJPDA, which update the states of objects using the posterior probability of joint associations, have high computational complexity due to the enumeration of joint associations. In [102], the number of measurements generated by an extended object is modeled as a truncated probability mass function, and a sum-product algorithm is developed to approximate the posterior of the objects iteratively.

The PMHT is investigated for extended object tracking in [61, 103]. In [104], a MEOT tracker based on PMHT for data assoication and random matrices [10] for extension estimation is proposed. The

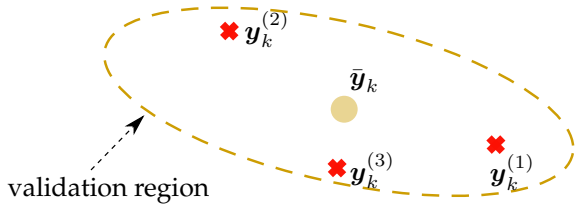


Figure 4.3: A PDA example. The predicted measurement mean $\bar{\boldsymbol{y}}_k$ is the orange dot. In this example, three measurements (red crosses) fall into the validation area (dashed ellipse).

PMHT is a batch method and normally is solved using iterative algorithms, such as the Expectation Maximization (EM) method.

4.2 Preliminaries

The multiple extended object tracker proposed in this chapter is based on the JIPDA filter, which is a single-scan and soft data association method. The background on JIPDA is introduced in this section. First, we start with the PDA measurement update for a single point object. Then, the JPDA filter, which deals with the problem that measurements having more than one possible origins caused by overlapping validation regions, is explained. The main idea of JIPDA is provided at the end of this section. The updated state and covariance after incorporating new measurements are differentiated by $^+$ signs in the following sections.

4.2.1 Probabilistic Data Association Filter

The PDA deals with the association ambiguities for the measurements which are located close to a predicted object (see Figure 4.3). To solve the association ambiguity, PDA evaluates the probability of each observation being a correct measurement for the object and performs a weighted update. In the meantime, the PDA filter increases the estimated error covariance to alleviate the effect of the possible erroneous assignment. The PDA filter is based on KF and assumes sufficient statistics have been obtained by incorporating measurements from previous time steps.

We restrict ourselves to parametric PDA, which assumes the number of false alarms is Poisson distributed with spatial density ρ. Based on the motion model, the predicted mean and covariance of the object are $\hat{\boldsymbol{x}}_k$ and $\mathbf{C}_k^{\boldsymbol{x}}$ at time step k. Using the measurement model in (3.3), the predicted measurement mean and covariance are obtained as $\bar{\boldsymbol{y}}_k$ and $\mathbf{C}_k^{\boldsymbol{y}}$, respectively. To reduce the complexity of data association, measurement validation or gating is often performed to restrict the number of valid association candidates. The region that satisfies

$$\mathcal{V}(\gamma) = \{\boldsymbol{y} : (\boldsymbol{y} - \bar{\boldsymbol{y}}_k)^{\mathrm{T}} \mathbf{C}_k^{\boldsymbol{y}} (\boldsymbol{y} - \bar{\boldsymbol{y}}_k) \leqslant \gamma\} \tag{4.1}$$

is called gating or validation region of the object with threshold γ. The probability that the true

measurement of the object falls in the region (4.1) is called gating probability P_G. It is often the case that multiple measurements fall in the validation region of one object (see Figure 4.3).

Denote the set of measurements in the validation region as $\mathbf{Y}_k = \{\boldsymbol{y}_k^{(1)}, \boldsymbol{y}_k^{(2)}, \cdots, \boldsymbol{y}_k^{(M_k)}\}$, where M_k is the number of measurements fall in the validation area at time step k. The PDA first calculates the predicted measurement likelihood for each validated measurement, i.e., $l_m = \mathcal{N}\left(\boldsymbol{y}_k^{(m)}; \bar{\boldsymbol{y}}_k, \mathbf{C}_k^{\boldsymbol{y}}\right)$. The probability of $\boldsymbol{y}_k^{(m)}$ being the correct measurement of the object is

$$\beta_m = \begin{cases} \frac{l_m P_D}{(1-P_D P_G)\rho + P_D \sum_{j=1}^{M_k} l_j} \ , & m = 1, \cdots, M_k \ , \\ \frac{1-P_D P_G}{(1-P_D P_G)\rho + P_D \sum_{j=1}^{M_k} l_j} \ , & m = 0 \ , \end{cases} \tag{4.2}$$

where P_D is the detection probability.

Then, PDA performs a weighted measurement update by incorporating all associations. Denoting $\tilde{\boldsymbol{y}}_k^{(m)} = \boldsymbol{y}_k^{(m)} - \bar{\boldsymbol{y}}_k$ as the innovation of $\boldsymbol{y}_k^{(m)}$, the object state is updated as

$$\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k + \mathbf{K}_k \sum_{m=1}^{M_k} \beta_m \tilde{\boldsymbol{y}}_k^{(m)} \ , \tag{4.3a}$$

where $\mathbf{K}_k = \mathbf{C}_k^{\boldsymbol{x}} \mathbf{H}_k^{\mathrm{T}} \left(\mathbf{C}_k^{\boldsymbol{y}}\right)^{-1}$ is the Kalman gain as explained in Section 3.2. The updated covariance consists of three parts:

- $\beta_0 \mathbf{C}_k^{\boldsymbol{x}}$, if none of the validated measurements is the true measurement;
- $(1-\beta_0)\left(\mathbf{C}_k^{\boldsymbol{x}} - \mathbf{K}_k \mathbf{C}_k^{\boldsymbol{y}} \mathbf{K}_k\right)$, with probability $(1-\beta_0)$, there is a correct measurement in $\mathbf{Y}_k$, then, the covariance is updated in the KF fashion but weighted with $(1-\beta_0)$;
- spread of innovations

$$\mathbf{K}_k \left(\sum_{m=1}^{M_k} \beta_m \tilde{\boldsymbol{y}}_k^{(m)} \left(\tilde{\boldsymbol{y}}_k^{(m)}\right)^{\mathrm{T}} - \sum_{m=1}^{M_k} \beta_m \tilde{\boldsymbol{y}}_k^{(m)} \left(\sum_{m=1}^{M_k} \beta_m \tilde{\boldsymbol{y}}_k^{(m)} \right)^{\mathrm{T}} \right) \mathbf{K}_k^{\mathrm{T}} \ , \tag{4.3b}$$

  the covariance is increased as it is not known which measurement is correct in $\mathbf{Y}_k$.

Combining these three parts, we have the updated covariance as

$$\begin{aligned} \mathbf{C}_k^{\boldsymbol{x}+} &= \beta_0 \mathbf{C}_k^{\boldsymbol{x}} + (1-\beta_0)\left(\mathbf{C}_k^{\boldsymbol{x}} - \mathbf{K}_k \mathbf{C}_k^{\boldsymbol{y}} \mathbf{K}_k^{\mathrm{T}}\right) \\ &+ \mathbf{K}_k \left(\sum_{m=1}^{M_k} \beta_m \tilde{\boldsymbol{y}}_k^{(m)} \left(\tilde{\boldsymbol{y}}_k^{(m)}\right)^{\mathrm{T}} - \sum_{m=1}^{M_k} \beta_m \tilde{\boldsymbol{y}}_k^{(m)} \left(\sum_{m=1}^{M_k} \beta_m \tilde{\boldsymbol{y}}_k^{(m)} \right)^{\mathrm{T}} \right) \mathbf{K}_k^{\mathrm{T}} \ . \end{aligned} \tag{4.3c}$$

If only one measurement $\boldsymbol{y}_k^{(1)}$ falls into the validation area, the probability of $\boldsymbol{y}_k^{(1)}$ originating from

the object is β_1. Then, the mean of the object is updated as

$$\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k + \beta_1 \mathbf{K}_k \tilde{\boldsymbol{y}}_k^{(1)} \quad . \tag{4.4a}$$

Applying (4.3c), the covariance is updated as [81]

$$\begin{aligned} \mathbf{C}_k^{\boldsymbol{x}+} &= (1-\beta_1)\mathbf{C}_k^{\boldsymbol{x}} + \beta_1 \left(\mathbf{C}_k^{\boldsymbol{x}} - \mathbf{K}_k \mathbf{C}_k^{\boldsymbol{y}} \mathbf{K}_k^{\mathrm{T}}\right) + \mathbf{K}_k \left(\beta_1 \tilde{\boldsymbol{y}}_k^{(1)} \left(\tilde{\boldsymbol{y}}_k^{(1)}\right)^{\mathrm{T}} - \beta_1^2 \tilde{\boldsymbol{y}}_k^{(1)} \left(\tilde{\boldsymbol{y}}_k^{(1)}\right)^{\mathrm{T}}\right) \mathbf{K}_k^{\mathrm{T}} \ , \\ &= \mathbf{C}_k^{\boldsymbol{x}} - \beta_1 \mathbf{K}_k \mathbf{C}_k^{\boldsymbol{y}} \mathbf{K}_k^{\mathrm{T}} + \beta_1 (1-\beta_1) \mathbf{K}_k \tilde{\boldsymbol{y}}_k^{(1)} \left(\mathbf{K}_k \tilde{\boldsymbol{y}}_k^{(1)}\right)^{\mathrm{T}} \quad . \end{aligned} \tag{4.4b}$$

4.2.2 Joint Probabilistic Data Association

In the case of overlapping gating regions caused by close objects, one observation can contribute to multiple trajectories (see Figure 4.4). The shared observations introduce a coupling into the measurement-object assignment because one measurement can be assigned to at most one object. The JPDA filter extends the PDA filter to a known number of targets and evaluates possible measurement-object assignments. After the removal of conflicting assignments and normalization for each track, a PDA update is performed independently on each track.

The JPDA filter assumes the number of objects is N_k at time k. Let θ_m be the object index to which measurement $\boldsymbol{y}_k^{(m)}$ is associated and row vector $\boldsymbol{\theta}_k = [\theta_1, \cdots, \theta_{M_k}]$ be the joint association variable for the new coming measurements set $\mathbf{Y}_k$. A valid joint association $\boldsymbol{\theta}_k$ satisfies the following constraints:

- each object can be assigned to at most one measurement which is located in its validation area,
- each measurement can be associated to at most one existing track.

Using the same clutter model as in Section 4.2.1, the probability of a joint association event is

$$P(\boldsymbol{\theta}_k \mid \mathbf{Y}_{1:k}) \propto \prod_{m=1}^{M_k} \rho^{-1} l_{m\theta_m} \prod_{\tau=1}^{N} (P_D)^{\delta_\tau} (1-P_D)^{1-\delta_\tau} \quad , \tag{4.5}$$

where $\mathbf{Y}_{1:k}$ is the set of measurements upto the k-th time step; $l_{m\tau}$ is the predicted likelihood, i.e., $\mathcal{N}\left(\boldsymbol{y}_k^{(m)}; \bar{\boldsymbol{y}}_k^{(\tau)}, \mathbf{C}_k^{\boldsymbol{y}(\tau)}\right)$; and δ_τ is binary and indicates the target detection status. Assuming the objects conditioned on the previous measurements are mutually independent, the state estimation is decoupled into N_k independent PDA updates. The marginal association probability $\beta_{m\tau}$ is obtained by summing up all the joint association events having measurement $\boldsymbol{y}_k^{(m)}$ assigned with object $\boldsymbol{x}_k^{(\tau)}$:

$$\beta_{m\tau} = \sum_{\boldsymbol{\theta}_k : \theta_m = \tau} P(\boldsymbol{\theta}_k \mid \mathbf{Y}_{1:k}) \quad . \tag{4.6}$$

After normalization, the state of each object are updated according to (4.3a) and (4.3c).

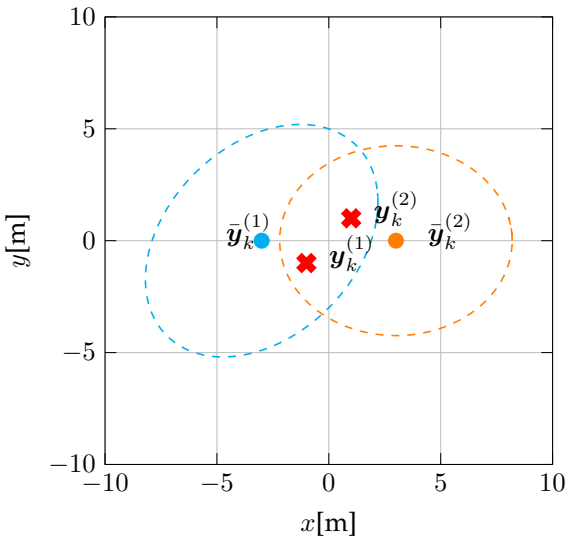


Figure 4.4: A JPDA example for point object tracking. This example gives two predicted objects with overlapping validation areas. Two predicted measurements are indicated as blue and orange dots. The validation areas (3σ contours) are dashed ellipses with respective colors. The measurements are crosses.

Similarly to the PDA filter, the JPDA filter uses a Gaussian distribution to represent the statistics from the past measurements. Weighted by the association probabilities, the exact statistic is a Gaussian mixture. Therefore, replacing the Gaussian mixture by a single Gaussian in (4.3a) and (4.3c) is an approximation. As we can see from Example 4.1, the calculation of $\beta_{m\tau}$ requires enumeration over all possible joint association events. The number of possible joint association events is combinatorial and yields an exponential complexity for JPDA.

4.2.3 Joint Integrated Probabilistic Data Association

As existing tracks might terminate and new tracks might enter the observation area, the JIPDA [80] extends JPDA by integrating an existence probability for each track. The evolvement of the existence probability of an object is modeled as a Markov chain. Two kinds of Markov chains are investigated in [85]. The Markov chain one [105, 85] assumes each track has two statuses: existent and non-existent. The Markov chain one is depicted in Figure 4.5. Assume the probability of an object τ being existent at times step k is $P_{\exists,k}^{(\tau)}$ and the probability of this object being non-existent is $P_{\nexists,k}^{(\tau)} = 1 - P_{\exists,k}^{(\tau)}$. The transition of the existence status can be described as

$$P_{\exists,k}^{(\tau)} = P_{11} P_{\exists,k-1}^{(\tau)} + P_{21}(1 - P_{\exists,k-1}^{(\tau)}) \ , \tag{4.7a}$$

$$P_{\nexists,k}^{(\tau)} = P_{12} P_{\exists,k-1}^{(\tau)} + P_{22}(1 - P_{\exists,k-1}^{(\tau)}) \ , \tag{4.7b}$$

Example 4.1. This example explains the calculation of marginal association probabilities for two point objects and two measurements as depicted in Figure 4.4 using JPDA. The predicted measurements are

$$\bar{\boldsymbol{y}}_k^{(1)} = \begin{bmatrix}-3, & 0\end{bmatrix}^{\mathrm{T}}, \quad \mathbf{C}_k^{\boldsymbol{y}(1)} = \begin{bmatrix}3 & 1\\ 1 & 3\end{bmatrix}, \quad \bar{\boldsymbol{y}}_k^{(2)} = \begin{bmatrix}3, & 0\end{bmatrix}^{\mathrm{T}}, \quad \mathbf{C}_k^{\boldsymbol{y}(2)} = \begin{bmatrix}3 & 0\\ 0 & 2\end{bmatrix}.$$

Two measurements are $\boldsymbol{y}_k^{(1)} = \begin{bmatrix}-1, & -1\end{bmatrix}^{\mathrm{T}}$ and $\boldsymbol{y}_k^{(2)} = \begin{bmatrix}1, & 1\end{bmatrix}^{\mathrm{T}}$. Assume we have the predicted measurement likelihoods as $l_{11} = 0.0172$, $l_{12} = 0.0035$, $l_{21} = 0.0038$, and $l_{22} = 0.0260$. In addition, detection probability is $P_D = 0.9$ and the density of clutter is $\rho = 0.001$. Each measurement has three possible origins: background (0), object 1, and object 2. With two measurements, there are up to 3^2 joint association possibilities. Among these nine joint associations, association $\boldsymbol{\theta}_k = [1, 1]$ and $\boldsymbol{\theta}_k = [2, 2]$ are invalid as one object is assigned with two measurements.
If $\boldsymbol{y}_k^{(1)}$ is clutter, the possible assignments and their probabilities (without normalization) are

$$\begin{array}{llllll}
 & \theta_2 = 0 & \boldsymbol{\theta}_{k,1}: & (1-P_D)^2\rho^2 & =1\times 10^{-8} & , \\
\theta_1 = 0 & \theta_2 = 1 & \boldsymbol{\theta}_{k,2}: & (1-P_D)\rho P_D l_{21} & =3.4463\times 10^{-7} & , \\
 & \theta_2 = 2 & \boldsymbol{\theta}_{k,3}: & (1-P_D)\rho P_D l_{22} & =2.3382\times 10^{-6} & .
\end{array}$$

If $\boldsymbol{y}_k^{(1)}$ is associated with object 1, similarly, we have

$$\begin{array}{llllll}
 & \theta_2 = 0 & \boldsymbol{\theta}_{k,4}: & P_D l_{11}(1-P_D)\rho & =1.5445\times 10^{-6} & , \\
\theta_1 = 1 & \theta_2 = 1 & - & - & - & , \\
 & \theta_2 = 2 & \boldsymbol{\theta}_{k,5}: & P_D^2 l_{11} l_{22} & =3.6114\times 10^{-4} & .
\end{array}$$

The joint associations where $\boldsymbol{y}_k^{(1)}$ is assigned with object 2 are:

$$\begin{array}{llllll}
 & \theta_2 = 0 & \boldsymbol{\theta}_{k,6}: & P_D l_{12}(1-P_D)\rho & =3.1644\times 10^{-7} & , \\
\theta_1 = 2 & \theta_2 = 1 & \boldsymbol{\theta}_{k,7}: & P_D^2 l_{12} l_{21} & =1.091\times 10^{-5} & , \\
 & \theta_2 = 2 & - & - & - & .
\end{array}$$

The Unnormalized marginal association probabilities for updating object 1 are

$$\begin{aligned}
\beta_{01} &= P(\boldsymbol{\theta}_{k,1} \mid \mathbf{Y}_{1:k}) + P(\boldsymbol{\theta}_{k,3} \mid \mathbf{Y}_{1:k}) + P(\boldsymbol{\theta}_{k,6} \mid \mathbf{Y}_{1:k}) && \propto 2.6645\times 10^{-6}\ , \\
\beta_{11} &= P(\boldsymbol{\theta}_{k,4} \mid \mathbf{Y}_{1:k}) + P(\boldsymbol{\theta}_{k,5} \mid \mathbf{Y}_{1:k}) && \propto 3.6269\times 10^{-4}\ , \\
\beta_{21} &= P(\boldsymbol{\theta}_{k,2} \mid \mathbf{Y}_{1:k}) + P(\boldsymbol{\theta}_{k,7} \mid \mathbf{Y}_{1:k}) && \propto 1.1250\times 10^{-5}\ .
\end{aligned}$$

After normalization, we have $\beta_{01} = 0.0071$, $\beta_{11} = 0.9630$, and $\beta_{21} = 0.0299$. The state and covariance of object 1 are updated by applying (4.3a) and (4.3c). The same process applies to object 2.

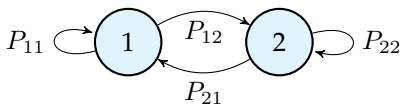


Figure 4.5: Markov chain one model for object existence. State 1 denotes the track exists while state 2 means the track is terminated. Transition probability from state i to state j is indicated using P_{ij}.

where P_{ij} for $i, j = 1, 2$ indicates the transition probability from state i to state j as described in Figure 4.5 The other Markov chain includes an extra status that an object exists but is not detected by the sensor, which will not be explained in detail in this section. The Markov chain one model is widely accepted in the tracking community. The transition probability P_{11} is also referred to as *survival probability* [32]. In addition, P_{21} most often is modeled as zero as a terminated track is assumed to be terminated forever. The existence probability for each track is updated on receiving the set of measurements and predicted to the next time step using (4.7a) in JIPDA.

The joint association variable $\boldsymbol{\theta}_k$ is defined as in JPDA. Let $\exists \boldsymbol{x}_k^{(\tau)}$ denote the event that object τ exists at time step k and $P_{\exists,k}^{(\tau)+}$ be the posterior existence probability of object τ given observation set $\mathbf{Y}_k$. The JIPDA filter first calculates the posterior probabilities of all possible association events, i.e., $P(\boldsymbol{\theta}_k \mid \mathbf{Y}_{1:k}), \forall \boldsymbol{\theta}_k$. The calculation of $P(\boldsymbol{\theta}_k \mid \mathbf{Y}_{1:k})$ is similar to (4.5). Then, the posterior probability of a marginal association event jointly with the existence of an object is obtained by summing the set of compatible association events, i.e.,

$$P\left(\theta_m = \tau, \exists \boldsymbol{x}_k^{(\tau)} \mid \mathbf{Y}_{1:k}\right) = \sum_{\boldsymbol{\theta}_k : \theta_m = \tau} P\left(\boldsymbol{\theta}_k \mid \mathbf{Y}_{1:k}\right) \quad . \tag{4.8a}$$

In addition, the existence probability of an object is updated by summing up the posterior probabilities of two sets of events: i) the object is existent, but no measurement is assigned and ii) the object is existent and is associated with one measurement:

$$P_{\exists,k}^{(\tau)+} = \sum_{\boldsymbol{\theta}_k : \forall m, \theta_m \neq \tau} P(\boldsymbol{\theta}_k, \exists\, \boldsymbol{x}_k^{(\tau)} \mid \mathbf{Y}_{1:k}) + \sum_{\boldsymbol{\theta}_k : \exists m, \theta_m = \tau} P(\boldsymbol{\theta}_k, \exists \boldsymbol{x}_k^{(\tau)} | \mathbf{Y}_{1:k}) \quad , \tag{4.8b}$$

where the posterior probability of object τ being existent but associated with no measurement is

$$\sum_{\boldsymbol{\theta}_k : \forall m, \theta_m \neq \tau} P(\boldsymbol{\theta}_k, \exists\, \boldsymbol{x}_k^{(\tau)} \mid \mathbf{Y}_{1:k}) = P_{\exists,k}^{(\tau)} \sum_{\boldsymbol{\theta}_k : \forall m, \theta_m \neq \tau} P(\boldsymbol{\theta}_k \mid \mathbf{Y}_{1:k}) \quad . \tag{4.8c}$$

In the last step, the conditional marginal association probability $\beta_{m\tau}$ is obtained as in [80]:

$$\beta_{m\tau} = \frac{P(\theta_m = \tau, \exists \boldsymbol{x}_k^{(\tau)} | \mathbf{Y}_{1:k})}{P_{\exists,k}^{(\tau)+}} \quad . \tag{4.8d}$$

The probability of object τ be associated with no measurements is

$$\beta_{0\tau} = \frac{\sum_{\boldsymbol{\theta}_k:\forall m,\theta_m\neq\tau} P(\boldsymbol{\theta}_k, \exists\, \boldsymbol{x}_k^{(\tau)} \mid \mathbf{Y}_{1:k})}{P_{\exists,k}^{(\tau)+}} \quad . \tag{4.8e}$$

The measurement update is conducted the same as in the JPDA and PDA filters.

4.3 Problem Description

At time step k, a set of measurements $\mathbf{Y}_k = \left\{\boldsymbol{y}_k^{(1)}, \cdots, \boldsymbol{y}_k^{(M_k)}\right\}$ is available. Each measurement $\boldsymbol{y}_k^{(m)} \in \mathbf{Y}_k$ originates from either an extended object or background. The false alarms, which are uniformly distributed on the observation area, are considered to be a Poisson point process with mean $\lambda_k^{(0)}$ and probability density l_0. The origins of the measurements are unknown.

Following the paradigm discussed in Chapter 3, an extended object gives rise to a varying number of error corrupted measurements. The object state is denoted as $\boldsymbol{x}_k^{(\tau)}$. The proposed tracker is based on the JIPDA filter and assumes the previous information up to time step $k-1$ is sufficiently summarized/approximated by state predictions, which are presented by mean and covariance of a Gaussian for each object. Let $\hat{\mathbf{X}}_k^{(0)} = \left\{\hat{\boldsymbol{x}}_k^{(\tau,0)}\right\}$ denote the set of objects that incorporates no measurements at time step k. The predicted set of objects $\hat{\mathbf{X}}_k^{(0)}$ consists of survived objects from the previous time step, possible birth components, and background. Without loss of generality, we assume the index of an extended object to be a positive natural number $\tau \in \mathbb{N}$. Index 0 is reserved for clutter. Let $\mathbf{N}_k$ be the set consisting of the labels in $\hat{\mathbf{X}}_k^{(0)}$. Similarly, $\mathbf{M}_k$ denotes the set of measurements' indices. The cardinalities of $\mathbf{N}_k$ and $\mathbf{M}_k$ are N_k and M_k, respectively.

The motion model of each object is assumed to be independent and follows the Markov chain process, which is described in Section 3.2. The number of tracked objects is varying over time. Existing objects might terminate, and new objects might appear. The set of extended objects $\mathbf{X}_k$ eliminates terminated tracks from time step $k-1$ and includes newborn objects at time step k.

On receiving the new measurements set $\mathbf{Y}_k$, the proposed multiple extended object tracker first incorporates a birth model and then conducts the measurement-object assignment. Based on the data association result, the MEOT is decoupled into several independent single EOT problems. As the single extended object tracker MEM-EKF* from Chapter 3 is used for each object, a sequential PDA algorithm is used to update the posterior probability of each object. After eliminating tracks with low existence probability, the persistent tracks are predicted to the next time step. The flow of the proposed tracker is shown in Algorithm 4.2 and Figure 4.6.

Algorithm 4.2 Structure of the proposed multiple extended object tracker MEM−JIPDA

1: **for** $k = 1, \cdots$ **do**
2: Calculate predicted likelihood (see Algorithm 4.4)
3: Perform data association
4: Track management (see Algorithm 4.5)
5: Multiple single-EOT using weighted MEM-EKF* (see Algorithm 4.6)
6: **end for**

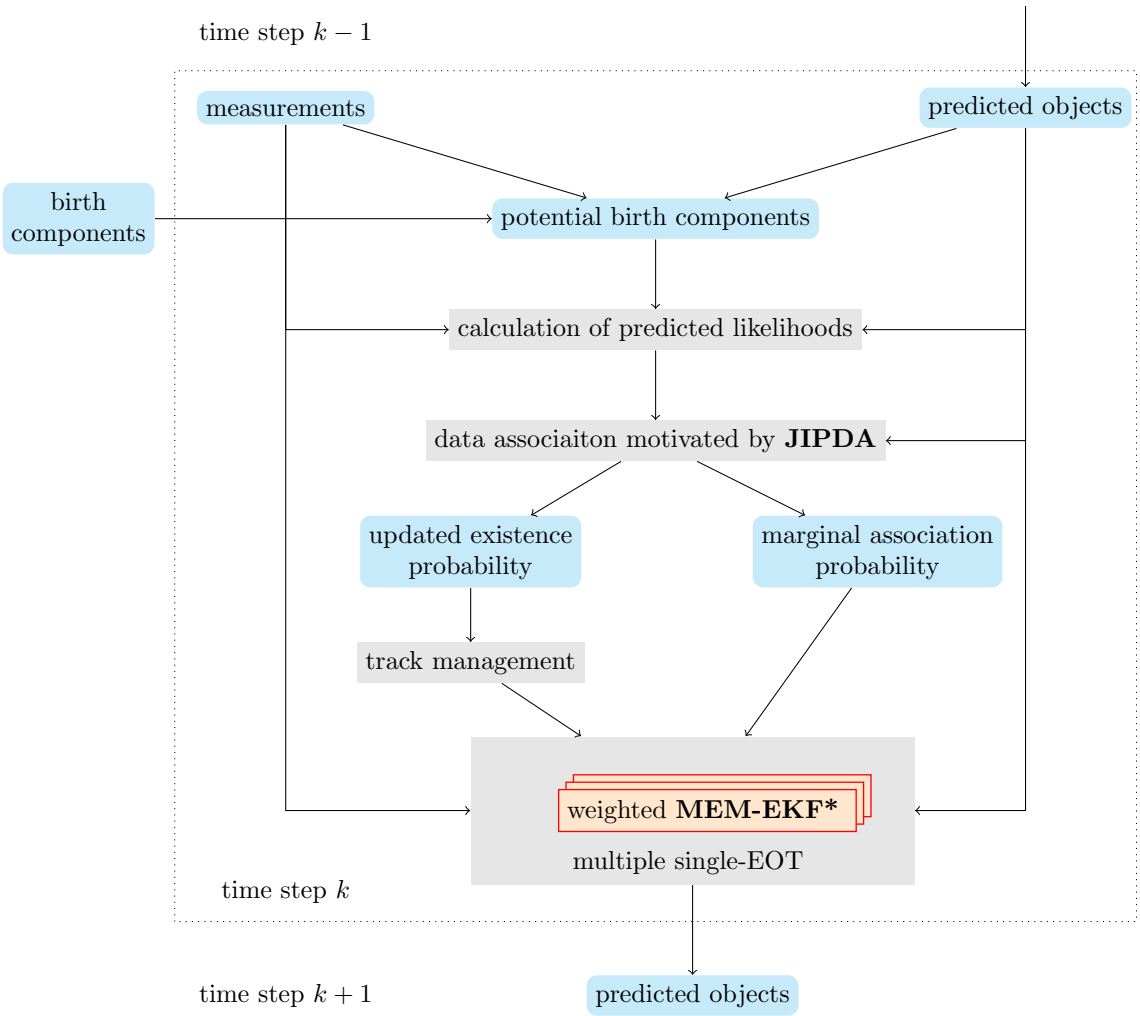


Figure 4.6: The structure of the MEM-JIPDA filter for multiple extended object tracking. The birth components are fixed for all time steps, but the potential birth components vary depending on the measurements and predicted objects. The track management process requires two thresholds: $\delta_{\text{terminate}}$ and δ_{initiate}, which determines the termination and initiation of tracks. Only the objects with significant existence probabilities are updated in the measurement update.

4.3.1 Assumptions and Approximations

The proposed MEOT tracker inherits all assumptions and approximations from the JPDA, JIPDA and PPP models. In addition, the state-dependent Poisson rates is approximated using the predicted states. This approximation is in accordance with the GIW-PHD filter [46]. All approximations and assumptions will be explained individually in the derivation steps.

4.3.2 Joint Association Hypothesis for Extended Objects

The definition of the joint association hypothesis is critical for JPDA based methods. As one object can be associated with multiple measurements, a measurement-oriented association hypothesis is employed. A joint association event that maps $\mathbf{M}_k$ to $\mathbf{N}_k$ is defined as:

$$\boldsymbol{\theta}_k = [\theta_1, \cdots, \theta_{M_k}] \ , \tag{4.9}$$

where $\forall m \in \mathbf{M}_k, \theta_m \in \mathbf{N}_k$ is the object label that $\boldsymbol{y}_k^{(m)}$ is associated with. [1] The space of associations from $\mathbf{M}_k$ to $\mathbf{N}_k$ is denoted as $\boldsymbol{\Theta}_{\mathbf{N}_k}^{\mathbf{M}_k}$. The structure of a joint association hypothesis defined in (4.9) ensures that a measurement can only be associated with at most one extended object. Furthermore, repetition is allowed for a single mapping in (4.9). This ensures the possibility that one object can get multiple measurements in a joint association event. A realization of $\boldsymbol{\theta}_k$ indicates a possible joint association event. Therefore, $\boldsymbol{\Theta}_{\mathbf{N}_k}^{\mathbf{M}_k}$ corresponds to $|\mathbf{N}_k|^{M_k}$ mutually exclusive and exhaustive joint association events. Let $\mathcal{Y}_k^{(\tau)}$ be the set of measurements and $n_k^{(\tau)}$ be the number of measurements that assigned to object τ. A joint association event is a partition on the set of measurements, which explicitly determines the number and the set of measurements that are assigned to each object, i.e., $\mathbf{T}_{\boldsymbol{\theta}_k} = \left\{ n_k^{(\tau)} \right\}_{\tau \in \mathbf{N}_k}$ and $\left\{ \mathcal{Y}_k^{(\tau)} \right\}_{\tau \in \mathbf{N}_k}$.

Example 4.2. Given three measurements $\mathbf{Y}_k = \{\boldsymbol{y}_k^{(1)}, \boldsymbol{y}_k^{(2)}, \boldsymbol{y}_k^{(3)}\}$ and predicted set of object: $\hat{\mathbf{X}}_k^{(0)} = \{\hat{\boldsymbol{x}}_0, \hat{\boldsymbol{x}}_k^{(1,0)}, \hat{\boldsymbol{x}}_k^{(2,0)}\}$, where $\hat{\boldsymbol{x}}_0$ indicates background. We have $\mathbf{N}_k = \{0, 1, 2\}$, $N_k = 3$, $\mathbf{M}_k = \{1, 2, 3\}$ and $M_k = 3$. Because each measurement has three possible origins, there are 27 joint associations. For example, $\boldsymbol{\theta}_k = [1, 0, 1]$ is one joint association hypothesis, which yields $\mathbf{T}_{\boldsymbol{\theta}_k} = \left\{ n_k^{(0)} = 1, n_k^{(1)} = 2, n_k^{(2)} = 0 \right\}$ and $\mathcal{Y}_k^{(0)} = \left\{ \boldsymbol{y}_k^{(2)} \right\}$, $\mathcal{Y}_k^{(1)} = \left\{ \boldsymbol{y}_k^{(1)}, \boldsymbol{y}_k^{(3)} \right\}$, $\mathcal{Y}_k^{(2)} = \emptyset$.

4.3.3 Zero-Inflated Poisson Point Process

The Inhomogeneous PPP model proposed in [18, 19] is used to model the observations produced by extended objects and background. The number of measurements $n_k^{(\tau)}$ generated by the extended

[1] Note that a joint association variable is bold while a single mapping is non-bold.

object τ is Poisson distributed with state-dependent rate parameter $\lambda_k^{(\tau)} := \lambda\left(\boldsymbol{x}_k^{(\tau)}\right)$, i.e.,

$$\text{Pois}\left(n_k^{(\tau)} \mid \lambda_k^{(\tau)}\right) = e^{-\lambda_k^{(\tau)}} \frac{\left(\lambda_k^{(\tau)}\right)^{n_k^{(\tau)}}}{n_k^{(\tau)}!} \quad . \tag{4.10}$$

The Poisson rate $\lambda_k^{(\tau)}$ can be designed as a function of the object area or a function of the distance between the object and the sensor. In addition, each measurement $\boldsymbol{y}_k^{(m)}$ generated by object τ is an independent random draw from the spatial distribution $p(\boldsymbol{y}_k^{(m)} \mid \boldsymbol{x}_k^{(\tau)})$.

As a higher Poisson rate implies that an object is likely to generate more measurements, the state-dependent Poisson rate can be seen as a built-in detection probability [18]. As described in [17], it cannot model complete occlusions. Therefore, the detection probability $P_{D,k}^{(\tau)}$ from traditional point object tracking is introduced additionally. This model is called *zero-inflated Poisson Point Process*. Due to this two-layer "detection probabilities", different from point object tracking, an extended object still has the possibility of generating no measurements when it is visible to the sensor. A so-called *effective probability of detection* is introduced to describe the probability that an extended object generates at least one measurement [17, 91] and is defined as $(1 - e^{-\lambda_k^{(\tau)}})P_{D,k}^{(\tau)}$. Given an object is detected, the counterpart of *effectively detected* is the case that an object is detected but generates no measurements ($P_{D,k}^{(\tau)} e^{-\lambda_k^{(\tau)}}$), i.e., *ineffectively detected*.

4.3.4 Existence Probability

The existence of each track is described by a so-called existence probability $P_{\exists,k}^{(\tau)}$. Combining detection and existence probability, Figure 4.7 describes the possible statuses of each object. Each track survives from time step k to time step $k+1$ with survival probability P_S. If the existence probability of a track is below a threshold $\delta_{\text{terminate}}$, the track is assumed to be terminated. A terminated track cannot be recovered to existence. Meanwhile, if the existence probability of a newborn component exceeds a threshold δ_{initiate}, a new track is considered to be born.

Remark 4.2. As the clutter takes part in the data association process as a background object, the clutter process is treated as an extended object and is included in $\mathbf{N}_k$ as object 0. The detection probability and existence probability are set as $P_{D,k}^{(0)} = 1$ and $P_{\exists,k}^{(0)} = 1$ for the seek of compact notation.

4.3.5 Predicted Measurement Likelihood

Each measurement in $\mathbf{Y}_k$ originates from an extended object or background. The clutter process is treated as a special object with index 0, the predicted measurement likelihood, given its associated

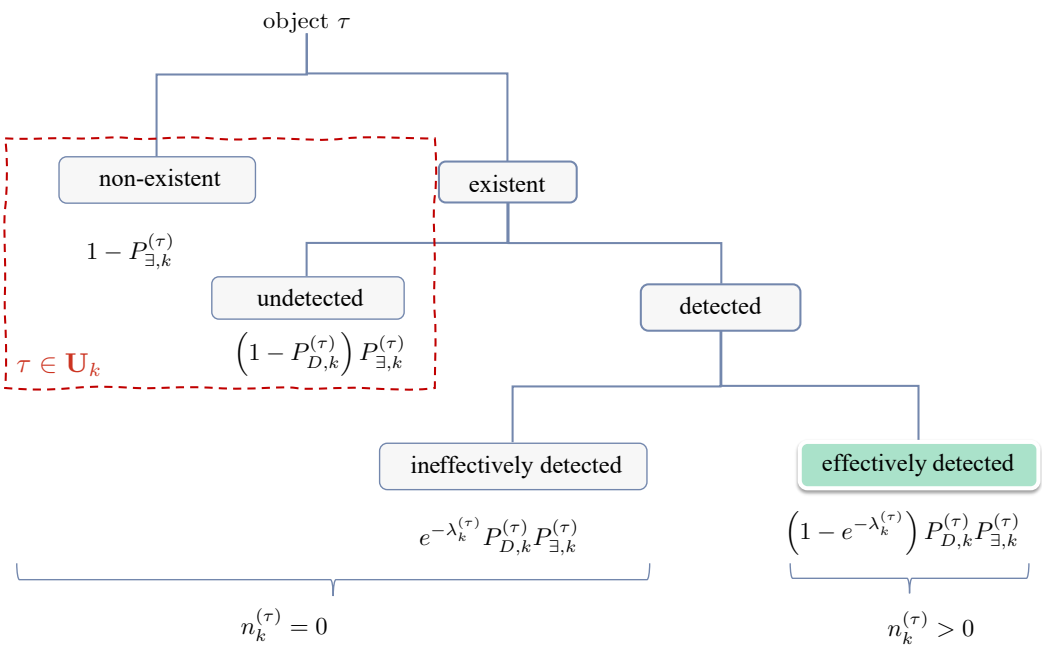


Figure 4.7: The possible existence and detection statuses of object τ. The number of assigned measurements $n_k^{(\tau)}$ is zero if object τ is non-existent, undetected or ineffectively detected. If object τ is either undetected or non-existent, we write $\tau \in \mathbf{U}_k$. [Yang3] ©2020IEEE

origin, is abbreviated as

$$p\left(\boldsymbol{y}_k^{(m)} \mid \theta_m, \mathbf{Y}_{1:k-1}\right) = \begin{cases} l_{m\theta_m}, \text{ if } \theta_m \in \mathbf{N}_k \ , \\ l_{m0}, \text{ if } \theta_m = 0 \ , \end{cases} \quad , \tag{4.11}$$

where $l_{m0} = l_0$ is the uniform probability density over the surveillance area. As the multiplicative error model presented in Chapter 3 is employed, the predicted single measurement likelihood is

$$l_{m\tau} = \mathcal{N}\left(\boldsymbol{y}_k^{(m)}; \bar{\boldsymbol{y}}_k^{(\tau,0)}, \mathbf{C}_k^{\boldsymbol{y}(\tau,0)}\right) \ , \tag{4.12}$$

where $\bar{\boldsymbol{y}}_k^{(\tau,0)}$ and $\mathbf{C}_k^{\boldsymbol{y}(\tau,0)}$ are the mean and covariance of the predicted measurement for object τ at time k according to (3.24) and (3.28). Given the predicted kinematic state $\hat{\boldsymbol{r}}_k^{(\tau,0)}$ and shape variable $\hat{\boldsymbol{p}}_k^{(\tau,0)}$ with covariance $\mathbf{C}_k^{\boldsymbol{r}(\tau,0)}$ and $\mathbf{C}_k^{\boldsymbol{p}(\tau,0)}$, the pseudo-code of the likelihood calculation is given in Algorithm 4.4.

4.4 State Prediction

The evolution of an existing object is modeled as a Markov process described in Section 3.2. In addition, all objects are assumed to move independently. To focus on the measurement update, a

unified linear motion model is used for all objects. If turning models are necessary, see Section 3.6.3 for modeling the correlation between orientation and turn rate. Based on the process model, traditional Kalman prediction applies for each existent object (see Section 3.2).

In addition to transition density, the existence status of each object needs to be considered in the prediction step. The existence probability for each track is assumed to be a Markov chain one [85] illustrated in Figure 4.5. Each track survives to the next time step with probability P_S. The prediction of the existence probability for each track follows

$$P_{\exists,k}^{(\tau)} = P_S P_{\exists,k-1}^{(\tau)+} \quad , \tag{4.13}$$

where $P_{\exists,k-1}^{(\tau)+}$ is the updated existence probability of object τ at time step $k-1$.

4.5 Data Association

This section first derives the posterior probability of a joint association event. By summing up the corresponding set of joint association events, the closed-form marginal association probabilities and existence probabilities are obtained similarly.

4.5.1 The Posterior Probability of a Joint Association Hypothesis

The key of JPDA based methods is to calculate the posterior probability of a joint association event. As measurements cannot be mapped to undetected and non-existent objects, the detection status of the extended objects is also considered. Denote the index-set of undetected and non-existent objects as $\mathbf{U}_k$, the set of detected objects' indices is $\mathbf{N}_k \backslash \mathbf{U}_k$.

A joint association $\boldsymbol{\theta}_k$ and a set of undetected and non-existent object indices $\mathbf{U}_k$ are compatible, if $\boldsymbol{\theta}_k$ does not map any measurements to objects in $\mathbf{U}_k$, i.e., $\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_{\mathbf{N}_k \backslash \mathbf{U}_k}^{\mathbf{M}_k}$. Applying Bayes rule, the joint posterior probability of an association hypothesis with the set of undetected objects is

$$\begin{aligned} P\left(\boldsymbol{\theta}_k, \mathbf{U}_k \mid \mathbf{Y}_{1:k}\right) &= P\left(\boldsymbol{\theta}_k, \mathbf{U}_k \mid \mathbf{Y}_k, \mathbf{Y}_{1:k-1}\right) \\ &\propto P\left(\mathbf{Y}_k \mid \boldsymbol{\theta}_k, \mathbf{U}_k, \mathbf{Y}_{1:k-1}\right) P\left(\boldsymbol{\theta}_k \mid \mathbf{U}_k, \mathbf{Y}_{1:k-1}\right) P\left(\mathbf{U}_k \mid \mathbf{Y}_{1:k-1}\right) \quad , \end{aligned} \tag{4.14}$$

if $\boldsymbol{\theta}_k$ and $\mathbf{U}_k$ are compatible, otherwise 0.

Example 4.3. Using the setting from Example 4.2, joint association $\boldsymbol{\theta}_k = [1, 0, 1]$ assigns no measurement to object 2. Therefore, $\mathbf{U}_k$ that is non-conflicting with $\boldsymbol{\theta}_k$ are $\{2\}$, which means object 2 is either terminated or undetected, and $\emptyset$, which indicates object 2 is ineffectively detected.

The three components in right hand side of (4.14) are derived. The JPDA assumes mutual indepen-

dence between objects given the measurements from previous time steps. In addition, the PPP model also assumes that measurements generated by each detected object are mutually independent. Enumerating all predicted objects (including clutter process) and measurements assigned to each object, we have the measurement likelihood conditioned on a joint association and a set of undetected/non-existent objects as

$$\begin{aligned} p\left(\mathbf{Y}_k \mid \boldsymbol{\theta}_k, \mathbf{U}_k, \mathbf{Y}_{1:k-1}\right) &= \prod_{\tau \in \mathbf{N}_k \backslash \mathbf{U}_k} \prod_{\boldsymbol{y}_k^{(m)} \in \mathcal{Y}_k^{(\tau)}} p\left(\boldsymbol{y}_k^{(m)} | \theta_m, \mathbf{Y}_{1:k-1}\right) \\ &= \prod_{\tau \in \mathbf{N}_k \backslash \mathbf{U}_k} \prod_{\boldsymbol{y}_k^{(m)} \in \mathcal{Y}_k^{(\tau)}} l_{m\tau} \ . \end{aligned} \tag{4.15}$$

Note that the ineffectively detected object indices are also included in $\mathbf{N}_k \backslash \mathbf{U}_k$. As no measurements are assigned to ineffectively detected objects, the latter product in (4.15) takes the value of one if the object is ineffectively detected.

In the second step, we derive the prior probability of a joint association variable given a set of undetected and non-existent objects. As a joint association event implicitly specifies the number of measurements assigned to each object, the prior of a joint association hypothesis conditioned on the set of undetected objects is

$$\begin{aligned} P(\boldsymbol{\theta}_k \mid \mathbf{U}_k, \mathbf{Y}_{1:k-1}) &= P(\boldsymbol{\theta}_k, \mathbf{T}_{\boldsymbol{\theta}_k} \mid \mathbf{U}_k, \mathbf{Y}_{1:k-1}) \\ &= P(\boldsymbol{\theta}_k \mid \mathbf{T}_{\boldsymbol{\theta}_k}, \mathbf{U}_k, \mathbf{Y}_{1:k-1}) P(\mathbf{T}_{\boldsymbol{\theta}_k} \mid \mathbf{U}_k, \mathbf{Y}_{1:k-1}) \ . \end{aligned} \tag{4.16a}$$

Given the numbers of measurements that are assigned to the objects, i.e., $\mathbf{T}_{\boldsymbol{\theta}_k}$, a joint association is assumed to be independent to $\mathbf{U}_k$ and the measurements from previous time steps, hence

$$P(\boldsymbol{\theta}_k \mid \mathbf{T}_{\boldsymbol{\theta}_k}, \mathbf{U}_k, \mathbf{Y}_{1:k-1}) = P(\boldsymbol{\theta}_k \mid \mathbf{T}_{\boldsymbol{\theta}_k}) \ . \tag{4.16b}$$

Joint association events $\boldsymbol{\theta}_k$s are a priori equally likely [106][107] given $\mathbf{T}_{\boldsymbol{\theta}_k}$. There are $\frac{M_k!}{\prod_{\tau \in \mathbf{N}_k} n_k^{(\tau)}}$ possible $\boldsymbol{\theta}_k$s based on combinatorics for a given $\mathbf{T}_{\boldsymbol{\theta}_k}$. We have

$$P(\boldsymbol{\theta}_k | \mathbf{T}_{\boldsymbol{\theta}_k}) = \frac{1}{M_k!} \prod_{\tau \in \mathbf{N}_k} n_k^{(\tau)}! \ . \tag{4.16c}$$

Assuming the numbers of observations generated by each object conditioned on all previous measurements are mutually independent, we have

$$P(\mathbf{T}_{\boldsymbol{\theta}_k} \mid \mathbf{U}_k, \mathbf{Y}_{1:k-1}) = \prod_{\tau \in \mathbf{U}_k} P\left(n_k^{(\tau)} \mid \mathbf{Y}_{1:k-1}\right) \prod_{\tau \in \mathbf{N}_k \backslash \mathbf{U}_k} P\left(n_k^{(\tau)} \mid \mathbf{Y}_{1:k-1}\right) \ . \tag{4.16d}$$

For an undetected and non-existent object, the number of measurements is 0, i.e., $P\left(0 \mid \mathbf{Y}_{1:k-1}\right) = 1$ for $\tau \in \mathbf{U}_k$. For detected objects, the number of measurements is Poisson distributed. Therefore,

the prior probability of $\mathbf{T}_{\boldsymbol{\theta}_k}$ in (4.16d) can be written as

$$P(\mathbf{T}_{\boldsymbol{\theta}_k} \mid \mathbf{U}_k, \mathbf{Y}_{1:k-1}) = \prod_{\tau \in \mathbf{N}_k \setminus \mathbf{U}_k} \text{Pois}\left(n_k^{(\tau)} \mid \lambda_k^{(\tau)}\right) \quad . \tag{4.16e}$$

As the real state of object τ is unknown, it is difficult to know the state-dependent true Poisson rate $\lambda_k^{(\tau)}$. The state dependent Poisson rates are approximated using the predicted object states, i.e., $\hat{\lambda}_k^{(\tau)} \approx \lambda\left(\hat{\boldsymbol{x}}_k^{(\tau)}\right)$. Then, the prior probability of $\mathbf{T}_{\boldsymbol{\theta}_k}$ is approximated as

$$P(\mathbf{T}_{\boldsymbol{\theta}_k} \mid \mathbf{Y}_{1:k-1}) \approx \prod_{\tau \in \mathbf{N}_k \setminus \mathbf{U}_k} \text{Pois}\left(n_k^{(\tau)} \mid \hat{\lambda}_k^{(\tau)}\right) \quad . \tag{4.16f}$$

Substituting (4.10) into (4.16f) and multiplying (4.16c) cancels out $\prod_{\tau \in \mathbf{N}_k \setminus \mathbf{U}_k} n_k^{(\tau)}!$ and gives us

$$P(\boldsymbol{\theta}_k \mid \mathbf{U}_k, \mathbf{Y}_{1:k-1}) = \prod_{\tau \in \mathbf{N}_k \setminus \mathbf{U}_k} \frac{e^{-\hat{\lambda}_k^{(\tau)}}}{M_k!} \left(\hat{\lambda}_k^{(\tau)}\right)^{n_k^{(\tau)}} \quad . \tag{4.16}$$

In the third step, we derive the prior probability of a set of undetected and non-existent objects. The prior probability of $\mathbf{U}_k$ is

$$P\left(\mathbf{U}_k \mid \mathbf{Y}_{1:k-1}\right) = \prod_{i \in \mathbf{U}_k} \left(1 - P_{D,k}^{(i)} P_{\exists,k}^{(i)}\right) \prod_{\tau \in \mathbf{N}_k \setminus \mathbf{U}_k} P_{D,k}^{(\tau)} P_{\exists,k}^{(\tau)} \quad . \tag{4.17}$$

Having (4.15), (4.16) and (4.17), (4.14) can be rewritten as

$$P\left(\boldsymbol{\theta}_k, \mathbf{U}_k \mid \mathbf{Y}_{1:k}\right) \propto \frac{1}{M_k!} \prod_{i \in \mathbf{U}_k} \left(1 - P_{D,k}^{(i)} P_{\exists,k}^{(i)}\right) \prod_{\tau \in \mathbf{N}_k \setminus \mathbf{U}_k} P_{D,k}^{(\tau)} P_{\exists,k}^{(\tau)} e^{-\hat{\lambda}_k^{(\tau)}} \left(\hat{\lambda}_k^{(\tau)}\right)^{n_k^{(\tau)}} \prod_{\boldsymbol{y}_k^{(m)} \in \mathcal{Y}_k^{(\tau)}} l_{m\tau} \quad . \tag{4.18a}$$

Let $c = \frac{1}{M_k!} \prod_{\tau \in \mathbf{N}_k} P_{D,k}^{(\tau)} P_{\exists,k}^{(\tau)} e^{-\hat{\lambda}_k^{(\tau)}}$ and $u_i = \frac{1 - P_{D,k}^{(i)} P_{\exists,k}^{(i)}}{P_{D,k}^{(i)} P_{\exists,k}^{(i)} e^{-\hat{\lambda}_k^{(i)}}}$, we have

$$P\left(\boldsymbol{\theta}_k, \mathbf{U}_k \mid \mathbf{Y}_{1:k}\right) \propto c \left(\prod_{i \in \mathbf{U}_k} u_i\right) \prod_{\tau \in \mathbf{N}_k \setminus \mathbf{U}_k} \left(\hat{\lambda}_k^{(\tau)}\right)^{n_k^{(\tau)}} \prod_{\boldsymbol{y}_k^{(m)} \in \mathcal{Y}_k^{(\tau)}} l_{m\tau} \quad . \tag{4.18b}$$

As the cardinality of $\mathcal{Y}_k^{(\tau)}$ is $n_k^{(\tau)}$, $\hat{\lambda}_k^{(\tau)}$ can be moved into the last product in (4.18b):

$$P\left(\boldsymbol{\theta}_k, \mathbf{U}_k \mid \mathbf{Y}_{1:k}\right) \propto \left(\prod_{i \in \mathbf{U}_k} u_i\right) \prod_{\tau \in \mathbf{N}_k \setminus \mathbf{U}_k} \prod_{\boldsymbol{y}_k^{(m)} \in \mathcal{Y}_k^{(\tau)}} \hat{\lambda}_k^{(\tau)} l_{m\tau} \quad . \tag{4.18c}$$

As the nested products in (4.18c) exhaust the measurements in $\mathbf{Y}_k$, it can be rewritten as an

enumeration of all measurements:

$$P\left(\boldsymbol{\theta}_k, \mathbf{U}_k \mid \mathbf{Y}_{1:k}\right) \propto \left(\prod_{i \in \mathbf{U}_k} u_i\right) \prod_{m \in \mathbf{M}_k} \hat{\lambda}_k^{(\theta_m)} l_{m\theta_m} \quad . \tag{4.18}$$

The posterior probability of an association jointly with a compatible $\mathbf{U}_k$ is given in (4.18). If $\boldsymbol{\theta}_k$ and $\mathbf{U}_k$ are conflicting, the probability is 0.

Remark 4.3. From (4.18), we can see that the mapping of a measurement is independent of the mappings of other measurements given the indices of undetected and non-existent objects. The probability of measurement $m \in \mathbf{M}_k$ being mapped to object $\tau \in \mathbf{N}_k \backslash \mathbf{U}_k$ is proportional to $\hat{\lambda}_k^{(\tau)} l_{m\tau}$. This is different from JPDA for point object tracking because a point object cannot be assigned to two measurements. Both the PPP model and the measurement-oriented joint association hypothesis are crucial for deriving this independence.

Remark 4.4. The posterior probability of an association hypothesis can be obtained by marginalization over $\mathbf{U}_k$, i.e.,

$$P\left(\boldsymbol{\theta}_k \mid \mathbf{Y}_{1:k}\right) \propto \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k)} \left(\prod_{i \in \mathbf{U}_k} u_i\right) \prod_{m \in \mathbf{M}_k} \hat{\lambda}_k^{(\theta_m)} l_{m\theta_m} \quad , \tag{4.19}$$

where $\mathbb{P}(\mathbf{N}_k)$ is the powerset of $\mathbf{N}_k$.

4.5.2 The Marginal Association Probability

In addition to the posterior probability of a joint association event, the estimation of the marginal association probability and existence probability relies on the following theorem.

Theorem 4.1. *Consider M ordered independent draws from set $\mathbf{D}$ with repetition, we denote a_{md} as a value that resulting from the m-th draw taking $d \in \mathbf{D}$. Denote $\boldsymbol{\theta} = \left[\theta_m\right]_{m=1}^{M}$ be a $1 \times M$ vector indicating a joint draw resulted by M draws, where $\theta_m \in \mathbf{D}$ is the result of M-th draw. Let $\boldsymbol{\Theta}_{\mathbf{D}}^{\mathbf{M}}$ be the set of all possible mappings from $\mathbf{M} = \{1, 2, \cdots, M\}$ to $\mathbf{D}$, We have*

$$\prod_{m \in \mathbf{M}} \left(\sum_{d \in \mathbf{D}} a_{md}\right) = \sum_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_{\mathbf{D}}^{\mathbf{M}}} \prod_{m=1}^{M} a_{m\theta_m} \quad . \tag{4.20}$$

Example 4.4. We have a set $\mathbf{D} = \{1, 2\}$, and want to draw $M = 2$ samples from set $\mathbf{D}$ such that the ordering matters and repetition is allowed. Let a_{md} be a value resulting from by the m-th draw taking $d \in \mathbf{N}$ (e.g., the probability of the m-th draw getting element d). The left hand side of (4.20) is

$$(a_{11} + a_{12})(a_{21} + a_{22}) \ .$$

There are 2^2 possible joint draws $\boldsymbol{\theta}$, i.e., $\boldsymbol{\Theta}_{\mathbf{D}}^{\mathbf{M}} = \{[1, 1], [1, 2], [2, 1], [2, 2]\}$. Substituting $\boldsymbol{\theta}$s to the right hand side of (4.20) gives us,

$$a_{11}a_{21} + a_{11}a_{22} + a_{12}a_{21} + a_{12}a_{22} \ ,$$

which is equal to the expanded left hand side of (4.20).

Proof. Theorem 4.1 can be proved using combinatorics or induction. The combinatorial proof is given here, and the induction proof is tedious and will not be given in this thesis. Without loss of generality, the set $\mathbf{D}$ is assumed to be $\{1, 2, \cdots, D\}$. The right hand side of (4.20) can be written as

$$(a_{11} + a_{12} + \cdots + a_{1D})(a_{21} + a_{22} + \cdots + a_{2D}) \cdots (a_{M1} + a_{M2} + \cdots + a_{MD}) \ . \tag{4.21}$$

Expanding (4.21) gives the sum of the D^M products of the form $a_{1_}a_{2_} \cdots a_{M_}$. It is the sum of all possible products obtained by taking one term from each factor. After (4.21) is expanded, each monomial of the form $a_{1_}a_{2_} \cdots a_{M_}$ corresponds to one joint draw, i.e., $\boldsymbol{\theta}$. The polynomial of expanding (4.21) is the exhaustion of all possible partitions. □

The right hand side of (4.20) has complexity $\mathcal{O}(|\mathbf{D}|^M)$ while the left hand side is $\mathcal{O}(M|\mathbf{D}|)$. An exponential summation is reduced to a linear product by applying Theorem 4.1.

To derive the exact formula for the marginal association probability, we enumerate the compatible $\mathbf{U}_k$s. According to Remark 4.3, given a set of undetected and non-existent objects, the mappings from the set of measurements to the set of detected objects are mutually independent, and hence Theorem 4.1 can be applied.

The marginalized probability of measurement m being associated with object τ is obtained by marginalizing over all compatible $\boldsymbol{\theta}_k$ and $\mathbf{U}_k$. Let $\mathbf{N}_k^\tau$ and $\mathbf{M}_k^m$ denote $\mathbf{N}_k \backslash \{\tau\}$ and $\mathbf{M}_k \backslash \{m\}$, respectively. Assigning object τ to measurement m implies that object τ is effectively detected. The compatible set of undetected and non-existent objects is $\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)$. The $\boldsymbol{\theta}_k$s that are compatible with $\theta_m = \tau$ and $\mathbf{U}_k$ are $\{\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_{\mathbf{N}_k \backslash \mathbf{U}_k}^{\mathbf{M}_k} : \theta_m = \tau\}$. Define $\boldsymbol{\theta}'_k$ as a $1 \times (M_k - 1)$ row vector indicating a joint mapping of $M_k - 1$ measurements. The set $\{\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_{\mathbf{N}_k \backslash \mathbf{U}_k}^{\mathbf{M}_k} : \theta_m = \tau\}$ can be reformulated as $\boldsymbol{\theta}'_k \in \boldsymbol{\Theta}_{\mathbf{N}_k \backslash \mathbf{U}_k}^{\mathbf{M}_k^m}$, which is a set of joint mappings from $\mathbf{M}_k^m$ to the detected object indices $\mathbf{N}_k \backslash \mathbf{U}_k$. According to the law of total probability, the marginal association probability of measurement m

being associated with object τ is

$$P(\theta_m = \tau \mid \mathbf{Y}_{1:k}) = \hat{\lambda}_k^{(\tau)} l_{m\tau} \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)} \sum_{\boldsymbol{\theta}'_k \in \boldsymbol{\Theta}_{\mathbf{N}_k \setminus \mathbf{U}_k}^{\mathbf{M}_k^m}} P(\boldsymbol{\theta}'_k, \mathbf{U}_k \mid \mathbf{Y}_{1:k}) \ . \tag{4.22a}$$

Substituting (4.18) into (4.22a) yields

$$P(\theta_m = \tau \mid \mathbf{Y}_{1:k}) \propto \hat{\lambda}_k^{(\tau)} l_{m\tau} \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}^\tau)} \left(\prod_{i \in \mathbf{U}_k} u_i \right) \sum_{\boldsymbol{\theta}'_k \in \boldsymbol{\Theta}_{\mathbf{N}_k \setminus \mathbf{U}_k}^{\mathbf{M}_k^m}} \prod_{j \in \mathbf{M}_k^m} \hat{\lambda}_k^{(\theta'_j)} l_{m\theta'_j} \ . \tag{4.22b}$$

Apply Theorem 4.1, we have the marginal association probability as

$$P(\theta_m = \tau \mid \mathbf{Y}_{1:k}) \propto \hat{\lambda}_k^{(\tau)} l_{m\tau} \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)} \left(\prod_{i \in \mathbf{U}_k} u_i \right) \prod_{g \in \mathbf{M}_k^m} \sum_{j \in \mathbf{N}_k \setminus \mathbf{U}_k} \hat{\lambda}_k^{(j)} l_{gj} \ . \tag{4.22}$$

As a measurement must have an origin, either an extended object or background clutter, the normalization is performed for each measurement. For each measurement $m \in \mathbf{M}_k$, we have

$$\sum_{\tau \in \mathbf{N}_k} P(\theta_m = \tau \mid \mathbf{Y}_{1:k}) = 1 \ . \tag{4.23}$$

Example 4.5. Using the same setting as in Example 4.2, this example gives a marginal association probability of measurement 1 being associated with object 1. Having $\mathbf{N}_k^1 = \{0, 2\}$ and $\mathbb{P}(\mathbf{N}_k^1) = \{\emptyset, \{0\}, \{2\}, \{0, 2\}\}$, the posterior probability of $\theta_1 = 1$ is

$$\begin{aligned} P(\theta_1 = 1 | \mathbf{Y}_{1:k}) \propto\ & \hat{\lambda}_k^{(1)} l_{11} \underbrace{(\hat{\lambda}_k^{(0)} l_{20} + \hat{\lambda}_k^{(1)} l_{21} + \hat{\lambda}_k^{(2)} l_{22})}_{\boldsymbol{y}_k^{(2)} \text{ has three possible mappings}} \underbrace{(\hat{\lambda}_k^{(0)} l_{30} + \hat{\lambda}_k^{(1)} l_{31} + \hat{\lambda}_k^{(2)} l_{32})}_{\boldsymbol{y}_k^{(3)} \text{ has three possible mappings}} \\ & + \hat{\lambda}_k^{(1)} l_{11} u_0 \underbrace{(\hat{\lambda}_k^{(1)} l_{21} + \hat{\lambda}_k^{(2)} l_{22})(\hat{\lambda}_k^{(1)} l_{31} + \hat{\lambda}_k^{(2)} l_{32})}_{\boldsymbol{y}_k^{(2)}, \boldsymbol{y}_k^{(3)} \text{ have two possible mappings given there is no clutter}} \\ & + \hat{\lambda}_k^{(1)} l_{11} u_2 (\hat{\lambda}_k^{(0)} l_{20} + \hat{\lambda}_k^{(1)} l_{21})(\hat{\lambda}_k^{(0)} l_{30} + \hat{\lambda}_k^{(1)} l_{31}) + \hat{\lambda}_k^{(1)} l_{11} u_0 u_2 \hat{\lambda}_k^{(1)} l_{21} \hat{\lambda}_k^{(1)} l_{31} \end{aligned}$$

4.5.3 Existence Probability Update

The existence probability and non-existence probability are calculated in a similar way as the marginalized association probability. In addition to the existence probability, the non-existence probability is also calculated in order to get the normalization factor.

Non-existence Probability

As object τ is specified as non-existent, we only consider the detection status of objects with indices in $\mathbf{N}_k^\tau$. As measurements must have at least one origin, $\mathbf{U}_k$ can only be a proper subset of $\mathbf{N}_k^\tau$. Denoting $\mathbb{P}(\mathbf{N}_k^\tau)^-$ as the powerset of $\mathbf{N}_k^\tau$ excluding the set $\mathbf{N}_k^\tau$ itself. The index-set of undetected and non-existent objects that is compatible with object τ being non-existent is $\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)^-$. The set of joint associations that are compatible with $\mathbf{U}_k$ and object τ being non-existent is $\boldsymbol{\theta}_k \in \boldsymbol{\Theta}^{\mathbf{M}_k}_{\mathbf{N}_k^\tau \setminus \mathbf{U}_k}$. Furthermore, u_τ needs to be modified as $u'_\tau = \frac{1-P^{(\tau)}_{\exists,k}}{e^{-\hat{\lambda}_k^{(\tau)}} P^{(\tau)}_{D,k} P^{(\tau)}_{\exists,k}}$ so that only the non-existence statuses are included. The posterior non-existence probability of object τ is

$$P^{(\tau)+}_{\nexists,k} = u'_\tau \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)^-} \sum_{\boldsymbol{\theta}_k \in \boldsymbol{\Theta}^{\mathbf{M}_k}_{\mathbf{N}_k^\tau \setminus \mathbf{U}_k}} P(\boldsymbol{\theta}_k, \mathbf{U}_k \mid \mathbf{Y}_{1:k}) \ . \tag{4.24a}$$

Substituting (4.18) gives us

$$P^{(\tau)+}_{\nexists,k} \propto u'_\tau \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)^-} \prod_{i \in \mathbf{U}_k} u_i \left(\sum_{\boldsymbol{\theta}_k \in \boldsymbol{\Theta}^{\mathbf{M}_k}_{\mathbf{N}_k^\tau \setminus \mathbf{U}_k}} \prod_{m \in \mathbf{M}_k} \hat{\lambda}_k^{(\theta_m)} l_{m\theta_m} \right) \ . \tag{4.24b}$$

Applying Theorem 4.1 gives us

$$P^{(\tau)+}_{\nexists,k} \propto u'_\tau \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)^-} \prod_{i \in \mathbf{U}_k} u_i \left(\prod_{m \in \mathbf{M}_k} \sum_{j \in \mathbf{N}_k^\tau \setminus \mathbf{U}_k} \hat{\lambda}_k^{(j)} l_{mj} \right) \ . \tag{4.24}$$

Existence probability

The existence probability of object τ consists of two cases: i) the probability that object τ exists but is not detected, and ii) the probability that object τ is detected.

First, we consider object τ being undetected but existent. Since object τ is undetected, it should not associate with any measurements. As $\mathbf{Y}_k$ must have at least one origin, $\mathbf{U}_k$ can only be a proper subset of $\mathbf{N}_k^\tau$. Therefore, the set of joint mappings that is compatible with the existence of object τ and $\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)^-$ is $\boldsymbol{\Theta}^{\mathbf{M}_k}_{\mathbf{N}_k^\tau \setminus \mathbf{U}_k}$ i.e., all joint mappings from $\mathbf{M}_k$ to $\mathbf{N}_k^\tau \setminus \mathbf{U}_k$. Similarly, the status that object τ is non-existent should be excluded from u_τ, i.e., $u_\tau - u'_\tau$. According to the law of total probability, the probability that τ exists but is undetected can be written as

$$(u_\tau - u'_\tau) \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)^-} \sum_{\boldsymbol{\theta}_k \in \boldsymbol{\Theta}^{\mathbf{M}_k}_{\mathbf{N}_k^\tau \setminus \mathbf{U}_k}} P(\boldsymbol{\theta}_k, \mathbf{U}_k \mid \mathbf{Y}_{1:k}) \ . \tag{4.25a}$$

Substituting (4.18) yields the posterior probability of object τ being undetected is

$$(u_\tau - u'_\tau) \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)^-} \prod_{i \in \mathbf{U}_k} u_i \left(\sum_{\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_{\mathbf{N}_k^\tau \setminus \mathbf{U}_k}^{\mathbf{M}_k}} \prod_{m \in \mathbf{M}_k} \hat{\lambda}_k^{(\theta_m)} l_{m\theta_m} \right) . \tag{4.25b}$$

Apply Theorem 4.1, we rewrite (4.25b) as

$$(u_\tau - u'_\tau) \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)^-} \prod_{i \in \mathbf{U}_k} u_i \left(\prod_{m \in \mathbf{M}_k} \sum_{j \in \mathbf{N}_k^\tau \setminus \mathbf{U}_k} \hat{\lambda}_k^{(j)} l_{mj} \right) . \tag{4.25c}$$

Let us consider the case that object τ is detected. The posterior probability that $\boldsymbol{x}_k^{(\tau)}$ is detected is the summation over $\mathbf{U}_k$ and its compatible $\boldsymbol{\theta}_k$s. If object τ is detected, the compatible $\mathbf{U}_k$ is a subset of $\mathbf{N}_k^\tau$. In this case, all objects with indices $\mathbf{N}_k^\tau$ can be undetected or non-existent. Because, if all objects with indices in $\mathbf{N}_k^\tau$ are undetected or non-existent, $\mathbf{Y}_k$ can be associated with object τ. The possible joint association events are $\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_{\mathbf{N}_k \setminus \mathbf{U}_k}^{\mathbf{M}_k}$. Hence, we have the posterior probability of object τ being detected as

$$\sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)} \sum_{\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_{\mathbf{N}_k \setminus \mathbf{U}_k}^{\mathbf{M}_k}} P(\boldsymbol{\theta}_k, \mathbf{U}_k \mid \mathbf{Y}_{1:k}) . \tag{4.25d}$$

Similarly, substituting (4.18) and applying Theorem 4.1 gives us the posterior probability of τ being detected as

$$\sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)} \prod_{i \in \mathbf{U}_k} u_i \left(\prod_{m \in \mathbf{M}_k} \sum_{j \in \mathbf{N}_k \setminus \mathbf{U}_k} \hat{\lambda}_k^{(j)} l_{mj} \right) . \tag{4.25e}$$

Summing up (4.25c) and (4.25e) gives us the updated existence probability of object τ as

$$\begin{aligned} P_{\exists,k}^{(\tau)+} \propto & \underbrace{(u_\tau - u'_\tau) \sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)^-} \prod_{i \in \mathbf{U}_k} u_i \left(\prod_{m \in \mathbf{M}_k} \sum_{j \in \mathbf{N}_k^\tau \setminus \mathbf{U}_k} \hat{\lambda}_k^{(j)} l_{mj} \right)}_{\text{object } \tau \text{ exists but is not detected}} \\ & + \underbrace{\sum_{\mathbf{U}_k \in \mathbb{P}(\mathbf{N}_k^\tau)} \prod_{i \in \mathbf{U}_k} u_i \left(\prod_{m \in \mathbf{M}_k} \sum_{j \in \mathbf{N}_k \setminus \mathbf{U}_k} \hat{\lambda}_k^{(j)} l_{mj} \right)}_{\text{object } \tau \text{ is detected}} . \end{aligned} \tag{4.25}$$

The normalization factor in (4.24) and (4.25) can be obtained by summing up (4.25) and (4.24) to unity.

4.5.4 Conditional Marginal Association Probability

As the existence of object τ is implied in the derivation of the marginal association probability in (4.22). The marginal association probability of measurement m is associated with object τ given the existence of object τ is

$$\beta_{m\tau} = \frac{P(\theta_m = \tau | \mathbf{Y}_{1:k})}{P_{\exists,k}^{(\tau)+}} \quad . \tag{4.26}$$

Since $\beta_{m\tau}$ is the marginal association probability conditioned on the existence of object τ, it is not necessary that $\sum_{\tau \in \mathbf{N}_k} \beta_{m\tau}$ equals to one. One only needs to ensure that the constraint described in (4.23) is fulfilled.

4.6 State Update

After the $\beta_{m\tau}$s have been determined, each target's state is updated in a PDA fashion by incorporating the association probabilities. The kinematics and shape variable update is based on the MEM-EKF* filter described in Chapter 3.

Since the measurement update of the MEM-EKF* filter is sequential, the measurements associated to object τ will be sequentially incorporated in the measurement update. Note that this sequential update is determined by the MEM-EKF* algorithm. If a measurement model updates the state of the extended object in a batch mode, such as the Random matrix approach, the marginal association probabilities should be incorporated correspondingly. At time step k, the predicted mean and covariance of the kinematic and shape variables for object τ are

$$\hat{\boldsymbol{r}}_k^{(\tau,0)}, \quad \mathbf{C}_k^{\boldsymbol{r}(\tau,0)} \text{ and } \hat{\boldsymbol{p}}_k^{(\tau,0)}, \quad \mathbf{C}_k^{\boldsymbol{p}(\tau,0)} \tag{4.27}$$

For each measurement $m \in \{1, \cdots, M_k\}$, if $\beta_{m\tau} > 0$, the cross-covariance $\mathbf{C}_k^{\boldsymbol{ry}(\tau,m)}$ and covariance $\mathbf{C}_k^{\boldsymbol{y}(\tau,m)}$ can be obtained using (3.27) and (3.28). Then, the Kalman gain is

$$\mathbf{K}_k^{\boldsymbol{r}(\tau,m)} = \mathbf{C}_k^{\boldsymbol{ry}(\tau,m)} \left(\mathbf{C}_k^{\boldsymbol{y}(\tau,m)} \right)^{-1} . \tag{4.28a}$$

The m-th mean of predicted measurement for object τ is

$$\bar{\boldsymbol{y}}_k^{(\tau,m)} = \mathbf{H}_k \hat{\boldsymbol{r}}_k^{(\tau,m-1)} \quad . \tag{4.28b}$$

The kinematic estimate of object τ is updated using measurement $\boldsymbol{y}_k^{(m)}$ and $\beta_{m\tau}$ as

$$\hat{\boldsymbol{r}}_k^{(\tau,m)} = \hat{\boldsymbol{r}}_k^{(\tau,m-1)} + \beta_{m\tau} \mathbf{K}_k^{\boldsymbol{r}(\tau,m)} \left(\boldsymbol{y}_k^{(m)} - \bar{\boldsymbol{y}}_k^{(\tau,m)} \right) \quad , \tag{4.28c}$$

which is an application of (4.4a).

For the covariance update, one modification is required compare to (4.4b). The directly application (4.4b) will introduce a biased estimation, especially for shape variables. This is because (4.4b) is a batch update for a point object while we are updating an extended object in a sequential manner. The coefficient for spread innovation is adjusted as the marginal association probability to the power of three has the best performance in simulations among some other intuitive parameters, such as $\frac{1}{M_k}$ and $\sum_m \beta_{m\tau}$. More investigation or justification of this modification will be explored in future work. The covariance is updated as

$$\begin{aligned}\mathbf{C}_k^{\boldsymbol{r}(\tau,m)} &= \mathbf{C}_k^{\boldsymbol{r}(\tau,m-1)} - \beta_{m\tau}\mathbf{K}_k^{\boldsymbol{r}(\tau,m)}\left(\mathbf{C}_k^{\boldsymbol{r}\boldsymbol{y}(\tau,m)}\right)^{\mathrm{T}} \\ &\quad + \beta_{m\tau}^3(1-\beta_{m\tau})\mathbf{K}_k^{\boldsymbol{r}(\tau,m)}\left(\boldsymbol{y}_k^{(m)} - \bar{\boldsymbol{y}}_k^{(\tau,m)}\right)\left(\boldsymbol{y}_k^{(m)} - \bar{\boldsymbol{y}}_k^{(\tau,m)}\right)^{\mathrm{T}}\left(\mathbf{K}_k^{\boldsymbol{r}(\tau,m)}\right)^{\mathrm{T}} \quad . \end{aligned} \tag{4.28d}$$

The update of the shape variable is similar to the kinematic state. First, the pseudo-measurement $\boldsymbol{\xi}_k^{(m)}$ and the m-th expected pseudo-measurement $\bar{\boldsymbol{\xi}}_k^{(\tau,m)}$ for object τ are obtained by applying (3.29) and (3.32). The construction of the pseudo-measurement is followed by the calculation of cross-covariance $\mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(\tau,m)}$ and covariance $\mathbf{C}_k^{\boldsymbol{\xi}(\tau,m)}$ using (3.34) and (3.33). The resulting Kalman gain of the shape variable $\hat{\boldsymbol{p}}_k^{(\tau,m-1)}$ given $\boldsymbol{\xi}_k^m$ is

$$\mathbf{K}_k^{\boldsymbol{p}(\tau,m)} = \mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(\tau,m)}\left(\mathbf{C}_k^{\boldsymbol{\xi}(\tau,m)}\right)^{-1} . \tag{4.29a}$$

Similar to the update of the kinematic state, the shape variable is updated as

$$\hat{\boldsymbol{p}}_k^{(\tau,m)} = \hat{\boldsymbol{p}}_k^{(\tau,m-1)} + \beta_{m\tau}\mathbf{K}_k^{\boldsymbol{p}(\tau,m)}\left(\boldsymbol{\xi}_k^{(m)} - \bar{\boldsymbol{\xi}}_k^{(\tau,m)}\right) \quad , \tag{4.29b}$$

$$\begin{aligned}\mathbf{C}_k^{\boldsymbol{p}(\tau,m)} &= \mathbf{C}_k^{\boldsymbol{p}(\tau,m-1)} - \beta_{m\tau}\mathbf{K}_k^{\boldsymbol{p}(\tau,m)}\left(\mathbf{C}_k^{\boldsymbol{p}\boldsymbol{\xi}(\tau,m)}\right) \\ &\quad + \beta_{m\tau}^3(1-\beta_{m\tau})\mathbf{K}_k^{\boldsymbol{p}(\tau,m)}\left(\boldsymbol{\xi}_k^{(m)} - \bar{\boldsymbol{\xi}}_k^{(\tau,m)}\right)\left(\boldsymbol{\xi}_k^{(m)} - \bar{\boldsymbol{\xi}}_k^{(\tau,m)}\right)^{\mathrm{T}}\left(\mathbf{K}_k^{\boldsymbol{p}(\tau,m)}\right)^{\mathrm{T}} \quad , \end{aligned} \tag{4.29c}$$

where $\beta_{m\tau}^3$ is used to balance the bias introduced by the sequential update in (4.29c) for extended objects. The sequential PDA update can be obtained by modifying the kinematic and shape update in Algorithm 3.1 (last four lines) using (4.28c),(4.28d), (4.29b) and (4.29c).

4.7 Implementation

This section presents the details of implementing the proposed MEOT tracker and its pseudo-code. The structure of the proposed tracker is described in Algorithm 4.2 and Figure 4.6. As the calculation of the conditioned marginal association probability and existence probability are given in (4.22) and (4.25), this section focuses on the birth model and track management. In addition, two suggestions on reducing the complexity are given at the end of this section.

4.7.1 Birth Model and Predicted Likelihoods

This section explains the calculation of the predicted likelihoods given the survived objects and potential birth components. The pseudo-code is given in Algorithm 4.4. Note that $\bar{\boldsymbol{y}}_k^{(\tau,1)}$ and $\mathbf{C}_B^{\boldsymbol{y}(\tau,1)}$ are the mean and covariance of predicted measurement of object τ at time k and can be obtained using Algorithm 4.3.

The estimate of an object is treated as a data structure, which contains

$$\hat{\boldsymbol{x}}_k^{(\tau,m)} = \left\{\hat{\boldsymbol{r}}_k^{(\tau,m)}, \hat{\boldsymbol{p}}_k^{(\tau,m)}, \mathbf{C}_k^{\boldsymbol{r}(\tau,m)}, \mathbf{C}_k^{\boldsymbol{p}(\tau,m)}, \hat{\lambda}_k^{(\tau,m)}, P_{\exists,k}^{(\tau,m)}, P_{D,k}^{(\tau,m)}\right\} \quad , \tag{4.30}$$

where the superscript m indicates that the estimate is obtained by updating with the m-th measurement. Correspondingly, the set of estimates is denoted as $\hat{\mathbf{X}}_k^{(m)} = \left\{\hat{\boldsymbol{x}}_k^{(\tau,m)}\right\}_{\tau\in\mathbf{N}_k}$. The set $\mathbf{N}_k$ consists of zero for the background, the indices of the survived tracks $\mathbf{N}_{S,k}$, and indices of potential birth components $\mathbf{N}_{B,k}$, i.e., $\mathbf{N}_k = \{0\} \cup \mathbf{N}_{S,k} \cup \mathbf{N}_{B,k}$. Note that the labeling algorithm is not the focus of this thesis, and we assume $\mathbf{N}_{S,k}$ and $\mathbf{N}_{B,k}$ are disjoint. In this subsection, the procedure of choosing potential birth components and predicted likelihoods are given.

Algorithm 4.3 Calculate the moments of predicted measurement for an extended object

Input:

- estimate of object τ: $\boldsymbol{x}_k^{(\tau,m)}$
- measurement noise covariance matrix of object τ: $\mathbf{C}_k^{\boldsymbol{v}(\tau)}$
- measurement matrix of object τ: $\mathbf{H}_k$

Output:

- predicted measurement mean and covariance: $\bar{\boldsymbol{y}}_k^{(\tau,m+1)}, \mathbf{C}_k^{\boldsymbol{y}(\tau,m+1)}$

1: $\bar{\boldsymbol{y}}_k^{(\tau,m+1)} = \mathbf{H}_k \hat{\boldsymbol{r}}_k^{(\tau,m)}$

2: $\begin{bmatrix}\alpha & l_1 & l_2\end{bmatrix}^{\mathrm{T}} = \hat{\boldsymbol{p}}_k^{(\tau,m)}$

3: $\mathbf{S} = \begin{bmatrix}\mathbf{S_1}\\ \mathbf{S_2}\end{bmatrix} = \begin{bmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{bmatrix}\begin{bmatrix}l_1 & 0\\ 0 & l_2\end{bmatrix}$

4: $\mathbf{J_1} = \begin{bmatrix}-l_1\sin\alpha & \cos\alpha & 0\\ -l_2\cos\alpha & 0 & -\sin\alpha\end{bmatrix}$

5: $\mathbf{J_2} = \begin{bmatrix}l_1\cos\alpha & \sin\alpha & 0\\ -l_2\sin\alpha & 0 & \cos\alpha\end{bmatrix}$

6: $\mathbf{C}^{\mathrm{I}} = \mathbf{S}\mathbf{C}^{\boldsymbol{h}}\mathbf{S}^{\mathrm{T}}$

7: $\mathbf{C}^{\mathrm{II}} = [\epsilon_{mn}] = \mathrm{tr}\left(\mathbf{C}_k^{\boldsymbol{p}^{(\tau,0)}}\mathbf{J}_n^{\mathrm{T}}\mathbf{C}^{\boldsymbol{h}}\mathbf{J}_m\right)$ for $m,n = 1,2$

8: $\mathbf{C}_k^{\boldsymbol{y}(\tau,m+1)} = \mathbf{H}_k\mathbf{C}_k^{\boldsymbol{r}(\tau)}\mathbf{H}_k^{\mathrm{T}} + \mathbf{C}^{\mathrm{I}} + \mathbf{C}^{\mathrm{II}} + \mathbf{C}_k^{\boldsymbol{v}(\tau)}$

The proposed tracker assumes fixed birth locations with respective uncertainties. These locations are places that are likely to have new objects entering the scene. Let $\hat{\mathbf{X}}_B$ denote the set of predefined birth components. Each birth components $\hat{\boldsymbol{x}}_B^{(\tau)} \in \hat{\mathbf{X}}_B$ contains the estimates of kinematic

Algorithm 4.4 Calculate predicted likelihoods

Input:

- measurement set: $\mathbf{Y}_k = \left\{ \boldsymbol{y}_k^{(m)} \right\}_{m=1}^{M_k}$
- set of birth components: $\hat{\mathbf{X}}_B$
- survived objects $\hat{\mathbf{X}}_{S,k}^{(0)}$ with set of indices $\mathbf{N}_{S,k}$
- gating threshold γ

Output:

- the set of survived objects and possible birth components $\hat{\mathbf{X}}_k^{(0)}$ with indices set $\mathbf{N}_k$
- predicted likelihoods: $l_{m\tau}, m = 1, \cdots, M_k, \tau \in \mathbf{N}_k$

$\mathbf{N}_k \leftarrow \mathbf{N}_{S,k}$
$\hat{\mathbf{X}}_k^{(0)} \leftarrow \hat{\mathbf{X}}_{S,k}^{(0)}$
for $\hat{\boldsymbol{x}}_k^{(\tau,0)} \in \hat{\mathbf{X}}_{S,k}^{(0)}$ **do**
    $\bar{\boldsymbol{y}}_k^{(\tau,1)}, \mathbf{C}_B^{\boldsymbol{y}(\tau,1)} \leftarrow$ calculate moments of predicted measurement for $\hat{\boldsymbol{x}}_k^{(\tau,0)}$ (Algorithm 4.3)
end for
for $m = 1, \cdots, M_k$ **do** ▷ compute predicted likelihoods
    for $\hat{\boldsymbol{x}}_k^{(\tau,0)} \in \hat{\mathbf{X}}_{S,k}^{(0)}$ **do**
        if $\left(\boldsymbol{y}_k^{(m)} - \bar{\boldsymbol{y}}_k^{(\tau,1)}\right)^{\mathrm{T}} \mathbf{C}_k^{\boldsymbol{y}(\tau,1)} \left(\boldsymbol{y}_k^{(m)} - \bar{\boldsymbol{y}}_k^{(\tau,1)}\right) \leqslant \gamma$ **then** ▷ gating
            $l_{m\tau} \leftarrow \mathcal{N}\left(\boldsymbol{y}_k^{(m)}; \bar{\boldsymbol{y}}_k^{(\tau,1)}, \mathbf{C}_k^{\boldsymbol{y}(\tau,1)}\right)$
        end if
    end for
    for $\hat{\boldsymbol{x}}_B^{(\tau)} \in \hat{\mathbf{X}}_B$ **do**
        if $\left(\boldsymbol{y}_k^{(m)} - \bar{\boldsymbol{y}}_B^{(\tau)}\right)^{\mathrm{T}} \mathbf{C}_B^{\boldsymbol{y}(\tau)} \left(\boldsymbol{y}_k^{(m)} - \bar{\boldsymbol{y}}_B^{(\tau)}\right) \leqslant \gamma$ **then** ▷ gating
            $l'_{m\tau} \leftarrow \mathcal{N}\left(\boldsymbol{y}_k^{(m)}; \bar{\boldsymbol{y}}_B^{(\tau)}, \mathbf{C}_B^{\boldsymbol{y}(\tau)}\right)$
        end if
        for $\hat{\boldsymbol{x}}_k^{(n)} \in \hat{\mathbf{X}}_{S,k}$ **do**
            if $l'_{m\tau} P_{\exists,B}^{(\tau)} > l_{mn} P_{\exists,k}^{(n)}$ **then** ▷ avoid the measurement competition
                $l_{mn} \leftarrow 0$
            else
                $l'_{m\tau} \leftarrow 0$
            end if
        end for
    end for
end for
for $\hat{\boldsymbol{x}}_B^{(\tau)} \in \hat{\mathbf{X}}_B$ **do**
    if $\exists m, l'_{m\tau} > 0$ **then**
        $\hat{\mathbf{X}}_k^{(0)} \leftarrow \hat{\mathbf{X}}_k^{(0)} \cup \{\hat{\boldsymbol{x}}_B^{(\tau)}\}$ ▷ append the potential birth components
        $\mathbf{N}_k \leftarrow \mathbf{N}_k \cup \{\tau\}$ ▷ expand the predicted likelihood matrix
    end if
end for
$\hat{\mathbf{X}}_k^{(0)} \leftarrow \hat{\mathbf{X}}_k^{(0)} \cup \{\hat{\boldsymbol{x}}_0\}$ ▷ append the background
$\mathbf{N}_k \leftarrow \mathbf{N}_k \cup \{0\}$
$l_{1:M_k,0} \leftarrow \rho$

state, shape variable, initial existence probability, estimated Poisson rate, and detection probability, the expected mean and covariance, which can be obtained using Algorithm 4.4, i.e., $\hat{\boldsymbol{x}}_B^{(\tau)} = \left\{\hat{\boldsymbol{r}}_B^{(\tau)}, \hat{\boldsymbol{p}}_B^{(\tau)}, \mathbf{C}_B^{\boldsymbol{r}(\tau)}, \mathbf{C}_B^{\boldsymbol{p}(\tau)}, P_{\exists,B}^{(\tau)}, \hat{\lambda}_k^{\tau}, P_{D,k}^{(\tau)}, \bar{\boldsymbol{y}}_B^{(\tau)}, \mathbf{C}_B^{\boldsymbol{y}(\tau)}\right\}$. Note that $\hat{\boldsymbol{x}}_B^{(\tau)}$ contains two more elements, predicted measurement mean and covariance, compared to a survived object's estimate (4.30). This is because the predicted measurement means and covariances for predicted objects are updated for each measurement while the birth components are fixed.

Before the data association step, a set of potential birth components is selected and then considered in the data association step. At scan k, only the birth components with measurement likelihoods greater than zero participate in the data association process (see line 13-15 in Algorithm 4.4). In other words, given measurement set $\mathbf{Y}_k$, if no measurement is located in the validation area of a birth component, this birth component will not be considered in the following data association and measurement update.

Furthermore, one must pay attention to the measurement competition between a birth component and a survived object. The measurement competition problem arises if a measurement lies in the overlapping validation area of a birth component and an existing track (most likely just confirmed its birth). To alleviate this problem, for each measurement m, its likelihood given a birth component is compared to the likelihood of measurement m given the survived object. Suppose the ratio is greater than the ratio of birth rate and corresponding existence probability. In this case, this measurement is assigned to the corresponding birth component instead of the survived object (see line 16-22 in Algorithm 4.4). This birth component with significant likelihoods is treated as an object and participates in the following data association step.

In conclusion, the confirmation of the birth of a new object must satisfy two conditions: first, it should have at least one measurement in its validation area with significant likelihoods; second, its updated existence probability after the data association should be greater than the initiation threshold.

Note that the likelihood of each measurement being clutter is appended in the last step. The clutter measurements are assumed to be uniformly distributed on the observation area, i.e., $l_{m0} = l_0$ with $m = 1, \cdots, M_k$. The background is denoted as $\hat{\boldsymbol{x}}^{(0)} = \left\{\hat{\lambda}_k^{(0)}, P_{\exists,k}^{(0)}, P_{D,k}^{(0)}\right\}$ which contains no kinematic and shape estimate, $P_{\exists,k}^{(0)} = 1$ and $P_{D,k}^{(0)} = 1$. [2]

4.7.2 Data Association

The conditional marginal association probabilities are implemented according to (4.26) and then normalized for each measurement because one measurement has only one origin. The existence probabilities are implemented using (4.25) and normalized with (4.24) for each object. Since the implementation of data association is straightforward, its pseudo-code is not repeated here. In

[2]Or some other values depending on the application. For example, $P_{\exists,k}^{(0)} = 1$ and $P_{D,k}^{(0)} < 1$, which allows the joint associations that assign no measurements as clutter.

cases of dense clutter, numerical underflow caused by taking the product over $\mathbf{M}_k$ in (4.22),(4.25) and (4.24) can be addressed using the log-sum-exp trick.

4.7.3 Track Management and Weighted MEM-EKF*

The guideline for track management is intuitive. If the updated existence probability of an existing track is below the termination threshold $\delta_{\text{termiante}}$, the track is considered as terminated at this scan. For a potential birth component, if its updated existence probability exceeds the initiation threshold δ_{initiate}, its birth is confirmed. Correspondingly, if the existence probability of a birth components is below δ_{initiate}, it is considered as a ghost track and removed from $\hat{\mathbf{X}}_k^{(0)}$ and $\mathbf{N}_k$. The pseudo-code of the track management is given in Algorithm 4.5.

Algorithm 4.5 Track Management

Input:
- the set of survived objects and possible birth components: $\hat{\mathbf{X}}_k^{(0)}$
- updated existence probability $P_{\exists,k}^{(\tau)+}, \tau \in \mathbf{N}_k \backslash \{0\}$

Output:
- the set of confirmed objects: $\hat{\mathbf{X}}_k^{(0)}$

```
for x̂_k^(τ,0) ∈ X̂_k^(0) do
    if τ ∈ N_{S,k} then
        if P_{∃,k}^(τ)+ < δ_terminate then
            τ is terminated, X̂_k^(0) \ {x̂_k^(τ,0)}
        end if
    else
        if P_{∃,k}^(τ)+ < δ_initiate then
            τ is a ghost track, X̂_k^(0) \ {x̂_k^(τ,0)}
        end if
    end if
end for
```

For each confirmed object, it is tracked using weighted MEM-EKF*. The pseudo-code is given in Algorithm 4.6. Note that the *sequential PDA upate* in Algorithm 4.6 (line 4) can be obtained by modifying the kinematic and shape update in Algorithm 3.1 (last four lines) using (4.28c),(4.28d), (4.29b) and (4.29c).

4.7.4 Discussion

The proposed data association step has complexity of $\mathcal{O}(M_k^2 N_k^2 2^{N_k})$. Obtaining a marginal association probability in (4.22) has a complexity of $\mathcal{O}(M_k N_k 2^{N_k})$, where $\mathcal{O}(2^{N_k})$ is for the calculation of the powerset of $\mathbf{N}_k$. As $M_k \times N_k$ marginal association probabilities need to be obtained, the overall complexity of the data association step is $\mathcal{O}(M_k^2 N_k^2 2^{N_k})$. As M_k and N_k is the number of

Algorithm 4.6 Measurement update and prediction using weighted MEM-EKF*

Input:

- confirmed objects: $\hat{\mathbf{X}}_k^{(0)}$
- measurement set: $\mathbf{Y}_k = \left\{ \boldsymbol{y}_k^{(m)} \right\}_{m=1}^{M_k}$
- conditioned marginal existence probability $\beta_{m\tau}, m = 1, \cdots, M_k, \tau \in \mathbf{N}_k \backslash \{0\}$

Output:

- the set of survived objects: $\hat{\mathbf{X}}_{k+1}^{(0)}$

$\hat{\mathbf{X}}_{k+1}^{(0)} = \emptyset$
for $\hat{\boldsymbol{x}}_k^{(\tau,0)} \in \hat{\mathbf{X}}_k^{(0)}$ **do**
  for $m = 1, \cdots, M_k$ **do** ▷ sequential PDA measurement update
    $\left\{ \hat{\boldsymbol{r}}_k^{(\tau,m)}, \hat{\boldsymbol{p}}_k^{(\tau,m)}, \mathbf{C}_k^{\boldsymbol{r}(\tau,m)}, \mathbf{C}_k^{\boldsymbol{p}(\tau,m)} \right\} \leftarrow$ according to Algorithm 3.1, (4.28c), (4.28d), (4.29b), and (4.29c)
  end for
  update $\hat{\boldsymbol{x}}_k^{(\tau,M_k)}$ using $\left\{ \hat{\boldsymbol{r}}_k^{(\tau,M_k)}, \hat{\boldsymbol{p}}_k^{(\tau,M_k)}, \mathbf{C}_k^{\boldsymbol{r}(\tau,M_k)}, \mathbf{C}_k^{\boldsymbol{p}(\tau,M_k)} \right\}$ and $P_{\exists,k}^{(\tau)+}$
  $\left\{ \hat{\boldsymbol{r}}_{k+1}^{(\tau,0)}, \hat{\boldsymbol{p}}_{k+1}^{(\tau,0)}, \mathbf{C}_{k+1}^{\boldsymbol{r}(\tau,0)}, \mathbf{C}_{k+1}^{\boldsymbol{p}(\tau,0)} \right\} \leftarrow$ prediction using motion model ▷ prediction
  predict existence probability $P_{\exists,k+1}^{(\tau)}$ using (4.13)
  update $\hat{\boldsymbol{x}}_{k+1}^{(\tau,0)}$ using $\left\{ \hat{\boldsymbol{r}}_{k+1}^{(\tau,0)}, \hat{\boldsymbol{p}}_k^{(\tau,0)}, \mathbf{C}_{k+1}^{\boldsymbol{r}(\tau,0)}, \mathbf{C}_k^{\boldsymbol{p}(\tau,0)} \right\}$ and $P_{\exists,k+1}^{(\tau)}$
  $\hat{\mathbf{X}}_{k+1}^{(0)} = \hat{\mathbf{X}}_{k+1}^{(0)} \cup \hat{\boldsymbol{x}}_{k+1}^{(\tau,0)}$
end for

measurements and objects that is considered in the data association process, the following two strategies can further reduce the complexity in the implementation:

1. only considering the potential birth components in the data association. That is to say, only the birth components satisfying: 1) have at least one measurement in their validation area, 2) the predicted likelihoods given birth components must be significant if birth components have overlapping gating area with the survived objects. This strategy is implemented in the birth model (see Algorithm 4.4).

2. dividing the predicted likelihoods, respective measurements, and predicted objects into small groups such that the objects across the different groups do not share measurements. Marginal association probability (4.22) and existence probability (4.25) considers all possible joint mappings between measurements and predicted objects. However, it is not necessary to jointly consider mappings for all objects if objects do not have overlapping validation areas. Therefore, the complexity of data association can be drastically reduced if the predicted likelihoods, respective measurements, and predicted objects are decomposed into groups such that objects from different groups do not share measurements. Then, the background $\hat{\boldsymbol{x}}^{(0)}$ is appended for each group. The data association and update are performed independently for each group. As this decoupling is intuitive, it is not explained in the pseudo-code for the seek of compactness.

4.8 Evaluation

In this section, the performance of the data association and MEOT tracker are evaluated. First, an example which gives an intuition on the proposed data association method is shown. Then, the proposed MEOT tracker is compared with a clustering and assignment based MEOT tracker called Gamma Gaussian Wishart Probability Hypothesis Density (GGIW-PHD)[92, 20] filter in a simulation. In the last experiment, the proposed tracker is evaluated using Lidar data to track pedestrians.

4.8.1 An Intuition on the Proposed Data Association Method

The first simulation gives an intuition on the proposed data association method. Two predicted objects and ten measurements are simulated. The simulation is illustrated in Figure 4.8. Two objects, object 1 (cyan) and object 2 (orange), with predicted existence probabilities of 0.9 are simulated. The existence probability for clutter is 1. Note that the cyan and orange dots and ellipses indicate the means and covariances (3σ contour) of the predicted measurements from the two respective objects, not the centers and extensions of the extended objects. For simplicity, the measurements are assumed to be Gaussian distributed. The detection probability for both predicted objects and clutter is 0.9. The simulated measurements are

$$\boldsymbol{y}_k^{(1)} = \begin{bmatrix} -19 \\ -19 \end{bmatrix}, \quad \boldsymbol{y}_k^{(2)} = \begin{bmatrix} 15 \\ -10 \end{bmatrix}, \quad \boldsymbol{y}_k^{(3)} = \begin{bmatrix} 10 \\ -15 \end{bmatrix}, \quad \boldsymbol{y}_k^{(4)} = \begin{bmatrix} -10 \\ 15 \end{bmatrix}, \quad \boldsymbol{y}_k^{(5)} = \begin{bmatrix} 0 \\ 10 \end{bmatrix},$$
$$\boldsymbol{y}_k^{(6)} = \begin{bmatrix} -9 \\ 5 \end{bmatrix}, \quad \boldsymbol{y}_k^{(7)} = \begin{bmatrix} -5 \\ 7 \end{bmatrix}, \quad \boldsymbol{y}_k^{(8)} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}, \quad \boldsymbol{y}_k^{(9)} = \begin{bmatrix} -1 \\ 6 \end{bmatrix}, \quad \boldsymbol{y}_k^{(10)} = \begin{bmatrix} 0 \\ -7 \end{bmatrix}.$$

The predicted measurements are

$$\bar{\boldsymbol{y}}_k^{(1,0)} = \begin{bmatrix} -4, & 5 \end{bmatrix}^{\mathrm{T}}, \quad \mathbf{C}_k^{\boldsymbol{y}(1,0)} = \operatorname{diag}\begin{bmatrix} 5, 2 \end{bmatrix},$$
$$\bar{\boldsymbol{y}}_k^{(2,0)} = \begin{bmatrix} -4, & -5 \end{bmatrix}^{\mathrm{T}}, \quad \mathbf{C}_k^{\boldsymbol{y}(2,0)} = \operatorname{diag}\begin{bmatrix} 7, 4 \end{bmatrix}.$$

In this simulation, the marginal association probabilities and existence probabilities of two objects are explained. As the predicted measurement likelihoods and Poisson rates are crucial for the proposed data association method, the sensitivity of the proposed data association method to the Poisson rates is tested for different Poisson rates.

An Intuition on the Proposed Data Association Method

The Possion rates for clutter and both predicted objects are 5. The marginal association probabilities and existence probabilities obtained using (4.22) and (4.25) are shown in Table 4.1. From Table 4.1, we can see that with a probability 0.9696 measurement 5 is a clutter while measurement 8 is a

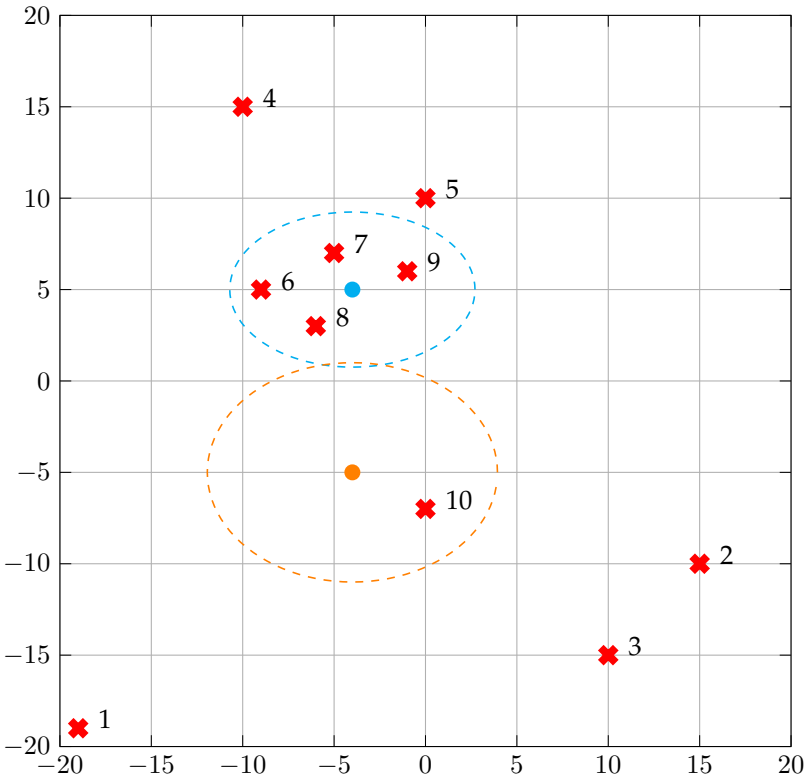


Figure 4.8: The predicted measurement means and covariances (3σ contours) for both predicted objects are plotted using dots and dashed ellipses, respectively. The red crosses are measurements, and their indices are labeled aside. [Yang3] ©2020IEEE

Table 4.1: The marginalized association probabilities obtained using the proposed data association technique for the example depicted in Figure 4.8. Note that the marginal association probabilities are not conditioned on the existence of the respective object. [Yang3] ©2020IEEE

	m	clutter	object 1	object 2
$P(\theta_m = \tau \mid \mathbf{Y}_{1:k})$	1	1	0	0
	2	1	0	0
	3	1	0	0
	4	1	0	0
	5	0.9696	0.0304	0
	6	0.1317	0.8683	0
	7	0.0362	0.9638	0
	8	0.0482	0.9518	0
	9	0.0380	0.9620	0
	10	0.7937	0	0.2063
$P_{\exists,k}^{(\tau)+}$	-	-	0.9998	0.5939

measurement from object 1 with probability 0.9518. With a probability of 0.2063, measurement 10 originates from object 2 while it is clutter with probability 0.7937. Note that the association probabilities shown in Table 4.1 are marginalized over the detection and existence status of object 2. Object 2 has only one measurement in its validation area, while it is expected to have five measurements if object 2 exists and is detected. Therefore, it is more likely measurement 10 is clutter, and object 2 is miss-detected or terminated.

As point object tracking assumes that at most one measurement is assigned to an object, the sum of the marginal probabilities for each object is one for traditional JPDA and JIPDA. However, the sum of the marginal probabilities for each object in the proposed data association is the expected number of measurements assigned to this object. For example, instead of one, the sum of marginal association possibilities of each object 1 is approximately 3.7763, which corresponds to the expected number of measurements object 1 gets at this scan. As the expected number of measurements assigned to object 1 is approximately 3.7763, and the Poisson rate of object 1 is five, the existence probability of object 1 increases from 0.9 to 0.9998. On the contrary, the existence probability of object 2 decreases to 0.5939 as only measurement 10 is assigned to object 2 with a probability of 0.2063.

Sensitivity to Poisson Rates

From (4.22) and (4.25), we know that the marginal association probability and existence probability depend on predicted measurement likelihoods and Poisson rates. As the state-dependent Poisson rates in (4.22) and (4.25) are approximated using the predicted states, using the same simulation setting depicted in Figure 4.8, it is investigated how marginal association probabilities and existence probabilities react to Poisson rates. The Poisson rates used in this evaluation for both objects are

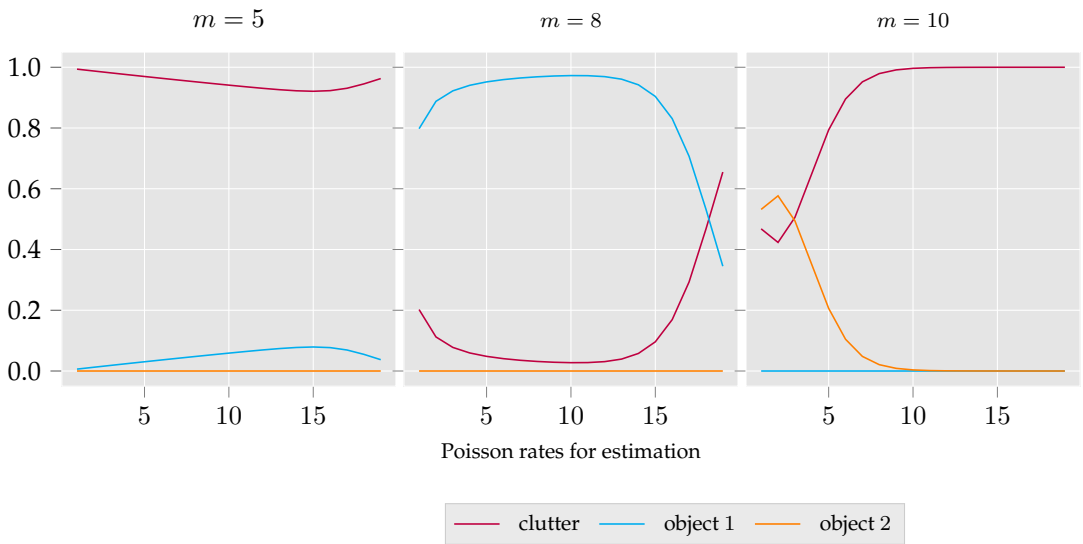


Figure 4.9: The estimated marginal probabilities using different Poisson rates, for example measurements 5, 8, and 10 in Figure 4.8. [Yang3] ©2020IEEE

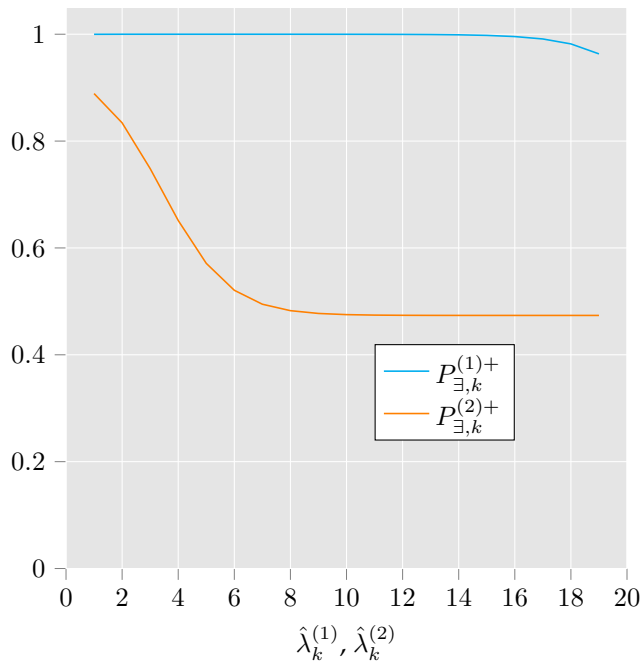


Figure 4.10: Updated existence probabilities using different Poisson rates for objects in Figure 4.8. [Yang3] ©2020IEEE

changed to values ranging from 1 to 20. Three example measurements, measurement 5, 8, and 10, are chosen to illustrate the obtained marginal association probabilities versus Poisson rates. The results are shown in Figure 4.9.

For a fixed set of measurements, an object tends to include more measurements as its Poisson rate increases. Therefore, we can see a slight increase in the probability that measurement 5 originates from object 1, which decreases the probability of measurement 5 being a clutter, correspondingly. The marginal probability of measurement 8 being a measurement of object 1 is almost one because its location is close to the predicted measurement mean of object 1. The probability of measurement 8 associated with object 1 is lower when the Poisson rate of object 1 is smaller than 3. This is because measurement 6, 7, and 9, which are also near the mean of predicted measurement, receives some probability. As the Poisson rate of object 1 becomes larger than 15, it is more likely that object 1 is undetected or non-existent given it has only 4 measurements near the mean of its predicted measurement. Correspondingly, the probability of measurement 6, 7, 8, and 9 being clutter becomes higher. Measurement 10 is on the edge of 3σ contour of object 2. Similarly, the probability of measurement 10 being a measurement of object 2 decreases as the estimated Poisson rate of object 2 increases, which indicates object 2 is more likely undetected or non-existent. In addition, Figure 4.10 depicts the change of existence probabilities of both objects against the Poisson rates.

4.8.2 Simulation with Closely-Spaced Objects

In this section, a simulation is constructed to assess the calculation of marginal association probabilities and existence probabilities. The MEM-JIPDA tracker is compared with GGIW-PHD filter [92, 20].

The GGIW-PHD filter models the extended objects as a Gamma Gaussian-inverse Wishart component. The Gamma distribution is the conjugate prior for the Poisson rate and the Gaussian-inverse Wishart distribution, which describes the measurements. A distance-based clustering and assignment method is employed to obtain the measurement partitions with significant weights. Then, a PHD filter is used to update the states of the extended objects. In addition, merging and pruning are required to manage the tracks.

We simulated two extended objects. The two objects move with a constant velocity of $13\,\mathrm{m/s}$ for 30 time steps. As two objects are approaching each other, they perform a coordinated turn with constant turn rates for 10 time steps to avoid a collision. They are then moving in parallel for 20 more steps and object one is terminated until the end of the simulation. The ground truth and simulated trajectories are shown in Figure 4.11. The goal of the first three motion phases of the simulation is to evaluate the estimation quality of the marginal association probabilities, as unsatisfactory marginal association probabilities will lead to poor state estimation in the first and second phases and track coalescence in the third phase. Object one is designed to terminate at time

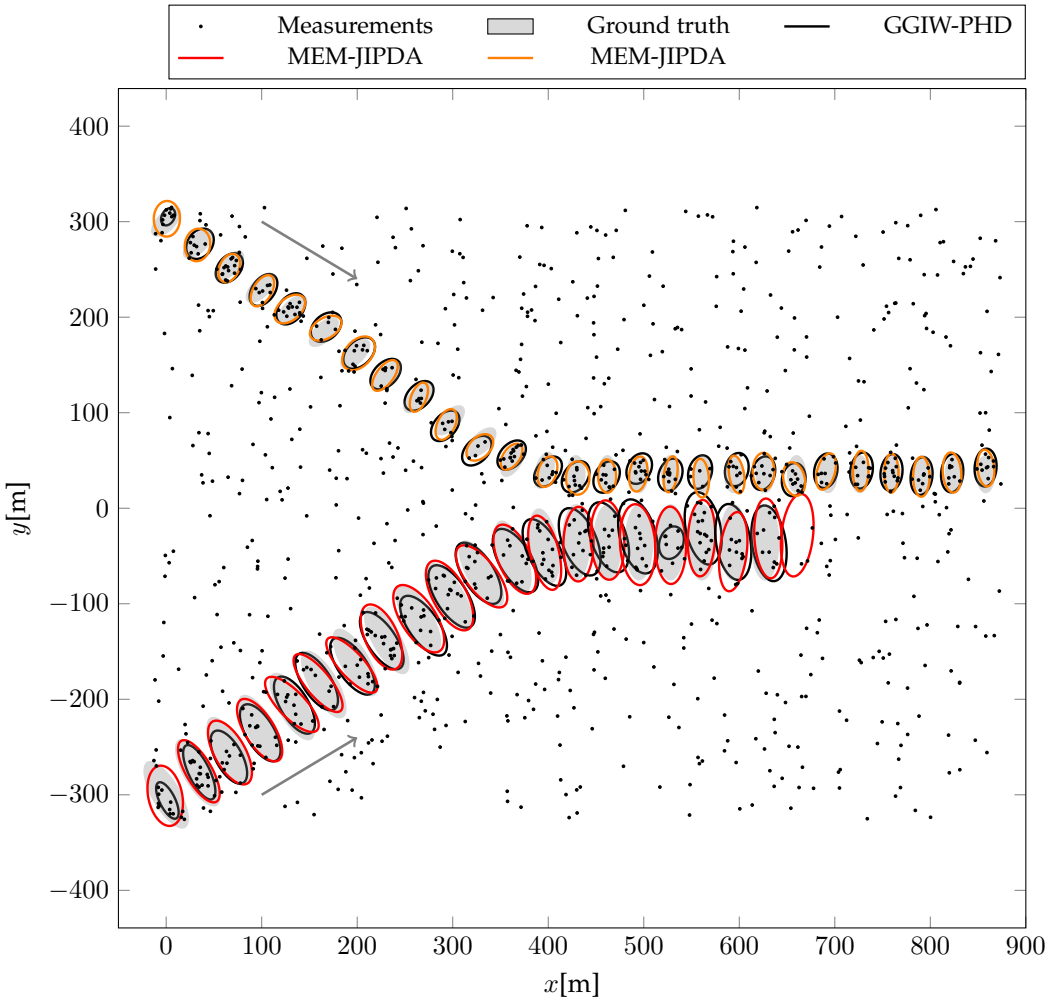


Figure 4.11: The simulated tracks, measurements, and one example run of the GGIW-PHD filter and the MEM-JIPDA filter. The grey arrows indicate the directions of two trajectories. As the original GGIW-PHD filter returns no label, its estimates are plotted using black ellipses. The two trajectories estimated using MEM-JIPDA are differentiated using orange and red. [Yang3] ©2020IEEE

step 60 to assess the existence probability estimation.

The number of clutter is Poisson distributed with mean 30, and the clutter is uniformly distributed across the observation area. The number of measurements that arise from each object is Poisson distributed with mean 12 and 9. The state vectors of the simulated objects initially are

$$\boldsymbol{r}_1^{(1,0)} = \begin{bmatrix} 0, & -300, & 11, & 7.7 \end{bmatrix}^{\mathrm{T}}, \qquad \boldsymbol{p}_1^{(1,0)} = \begin{bmatrix} -\frac{\pi}{3}, & 40, & 15 \end{bmatrix}^{\mathrm{T}},$$
$$\boldsymbol{r}_1^{(2,0)} = \begin{bmatrix} 0, & -300, & 11, & -7.7 \end{bmatrix}^{\mathrm{T}}, \qquad \boldsymbol{p}_1^{(2,0)} = \begin{bmatrix} \frac{\pi}{4}, & 20, & 10 \end{bmatrix}^{\mathrm{T}}.$$

The priors for both objects are

$$\hat{\boldsymbol{r}}_1^{(1,0)} = \begin{bmatrix} -10, & -280, & 10, & 10 \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{C}_1^{\boldsymbol{r}(1,0)} = \mathrm{diag}\begin{bmatrix} 200, & 200, & 10, & 10 \end{bmatrix},$$
$$\hat{\boldsymbol{r}}_1^{(2,0)} = \begin{bmatrix} -10, & 280, & 10, & -10 \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{C}_1^{\boldsymbol{r}(2,0)} = \mathrm{diag}\begin{bmatrix} 200, & 200, & 10, & 10 \end{bmatrix},$$
$$\hat{\boldsymbol{p}}_1^{(1,0)} = \begin{bmatrix} 0, & 20, & 20 \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{C}_1^{\boldsymbol{p}(1,0)} = \mathrm{diag}\begin{bmatrix} 0.05, & 40, & 40 \end{bmatrix},$$
$$\hat{\boldsymbol{p}}_1^{(2,0)} = \begin{bmatrix} 0, & 20, & 20 \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{C}_1^{\boldsymbol{p}(2,0)} = \mathrm{diag}\begin{bmatrix} 0.05, & 40, & 40 \end{bmatrix}.$$

The Poisson rates used in the estimation are the same as the true Poisson rates. The initial existence probability for both objects is 0.9. The survival probability for both tracks is 0.8, and the detection probability is 0.9. Two birth components are specified as

$$\hat{\boldsymbol{r}}_B^{(1)} = \begin{bmatrix} -20, & -250, & 10, & 10 \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{C}_B^{\boldsymbol{r}(1)} = \mathrm{diag}\begin{bmatrix} 900, & 900, & 10, & 10 \end{bmatrix},$$
$$\hat{\boldsymbol{r}}_B^{(2)} = \begin{bmatrix} -20, & 250, & 10, & -10 \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{C}_B^{\boldsymbol{r}(2)} = \mathrm{diag}\begin{bmatrix} 900, & 900, & 10, & 10 \end{bmatrix},$$
$$\hat{\boldsymbol{p}}_B^{(1)} = \begin{bmatrix} 0, & 30, & 30 \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{C}_B^{\boldsymbol{p}(1)} = \mathrm{diag}\begin{bmatrix} 0.1, & 40, & 40 \end{bmatrix},$$
$$\hat{\boldsymbol{p}}_B^{(2)} = \begin{bmatrix} 0, & 15, & 15 \end{bmatrix}^{\mathrm{T}}, \qquad \mathbf{C}_B^{\boldsymbol{p}(2)} = \mathrm{diag}\begin{bmatrix} 0.1, & 40, & 40 \end{bmatrix}.$$

The initial existence probability for both components is 0.1, and their Poisson rates are 10. If the existence probability of a new component exceeds 0.8, we confirm its birth. When the existence probability of an exiting track falls below 0.2, the track is treated as terminated.

One comparison example run is illustrated in Figure 4.11. In addition, the OSPA distances [30] combined with the Gaussian Wasserstein distance [Yang9] with a cut off value of 100 for 100 Monte Carlo runs are calculated for comparison. Besides the mean OSPA distance, the one standard deviation are also shown in Figure 4.12. From the figure, we can see that MEM-JIPDA has a smaller error in terms of mean and variance overall, except that MEM-JIPDA spiked at the time step 61. This is because object 1 terminates at time 61, but the MEM-JIPDA considers it as miss-detected. MEM-JIPDA has better performance, in terms of mean and standard deviation, for the rest of scans.

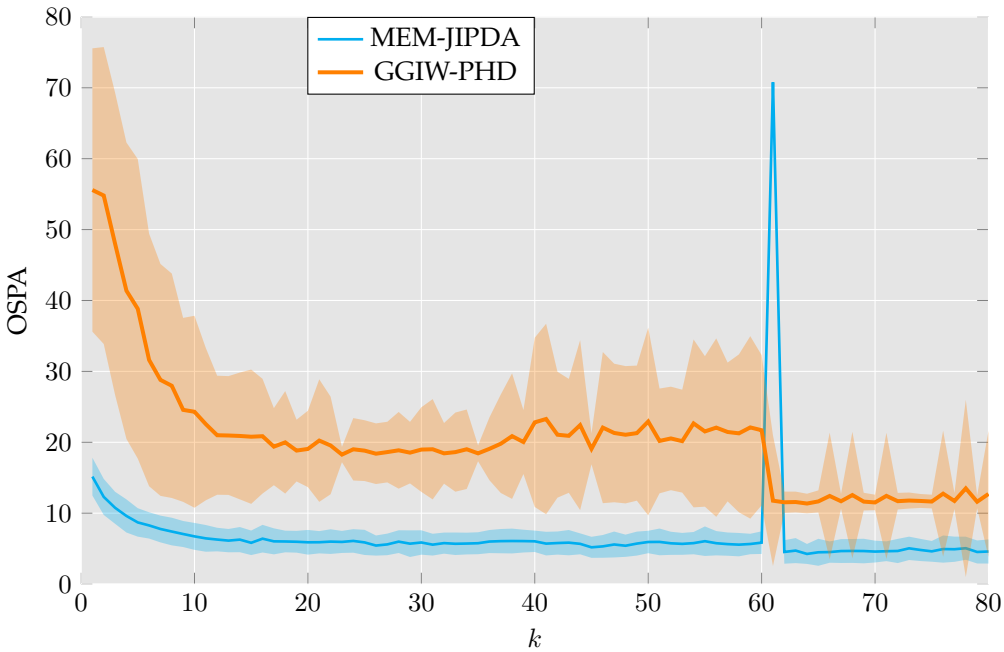


Figure 4.12: Mean OSPA distance ± one standard deviation for the simulation with two closely-spaced objects. [Yang3] ©2020IEEE

Note that, compared to MEM-JIPDA, the GGIW-PHD filter uses a spontaneous birth model, which is more flexible than MEM-JIPDA and requires less information about the observation area. On the other hand, if multiple clutter measurements form a cluster, the GGIW-PHD filter could interpret this clutter-cluster as a newborn object.

4.8.3 Tracking Pedestrians using Lidar Data

Furthermore, the proposed tracker is tested for tracking lidar data from pedestrians. The lidar data is from the KITTI dataset [2]. The KITTI dataset is a well-known benchmark for performing pedestrian and vehicle detection or tracking tasks. It consists of many scenarios, such as vehicles on the street, people walking in residential areas, and students on the campus. Multiple sensors, such as GPS/IMU, Laserscanner, Grayscale camera, and RGB camera, are mounted on a car. The scenes are recorded while the car is driving or parking. The scene applied in this experiment is from the dataset "`2011_09_28_drive_0043`".[3] This scene is recorded on the campus of Karlsruhe Institute of Technology. The scene, sensor setup and one example lidar data is shown in Figure 1.2 in Chapter 1.

A rectangular area, i.e., $[250, \quad 750]\mathrm{cm} \times [-500, \quad 150]\mathrm{cm}$ is trimmed to be the observation area. The tracking algorithm is performed on the first 50 time steps. The lidar points, approximately

[3] `http://www.cvlibs.net/datasets/kitti/raw_data.php?type=campus`

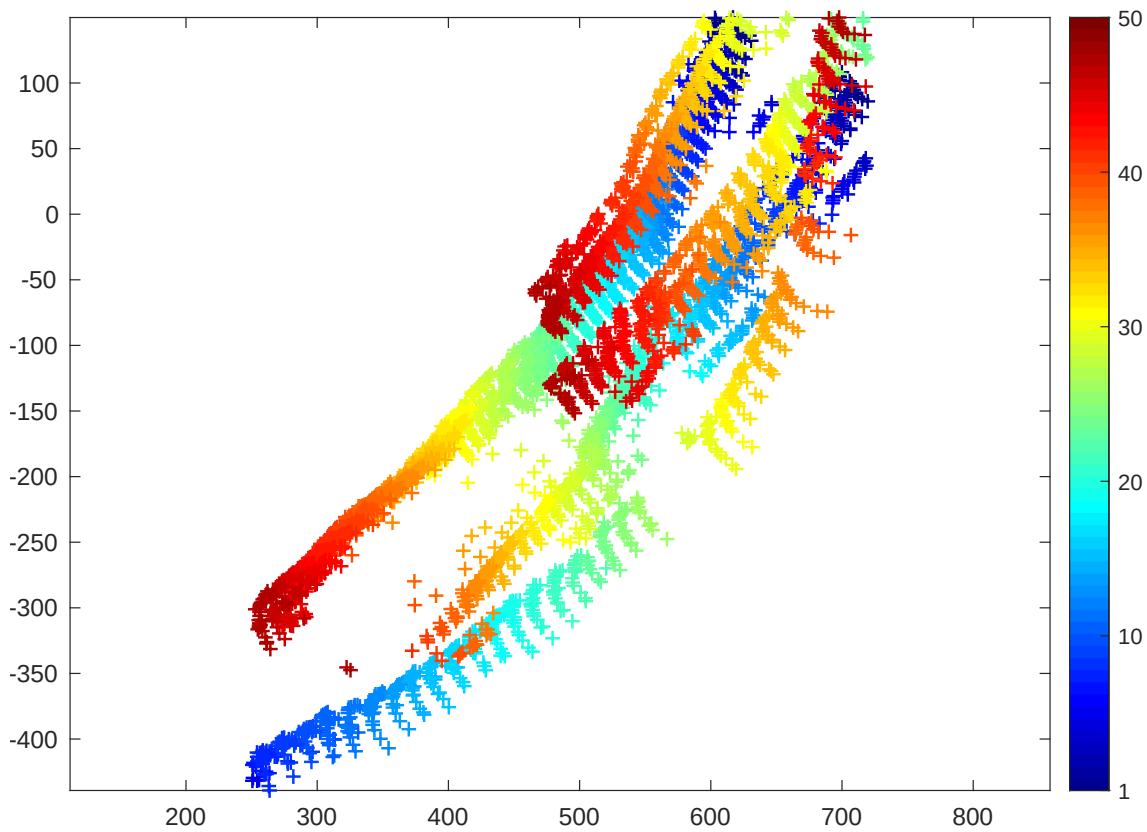


Figure 4.13: This figure depicts the KITTI pedestrian lidar measurements over time. Measurements available at different time steps are shown in different colors. The color bar on the right indicates the corresponding time steps. Note that two times of occlusions of one person can be seen in this figure. [Yang3] ©2020IEEE

the height of knees, at the height of 39-43 cm, are selected as measurements. Hence, the shapes are approximated as ellipses. In the first 50 time steps, five people appeared in the tracking area. Person 1 and 2 appear at the first time step. Person 2 is occluded by person 1 from the scan 45 until the end. Person 3 enters the observation area at time step 6 and is completely occluded at time step 29, 30, 40 and 43. The overall data is shown in Figure 4.13, where the color indicates the time step.

Parameters

As a proper birth model is assumed to detect the appearance of objects, no prior is given in this experiments. The four birth locations in this experiment are

$$\begin{aligned}
\hat{\boldsymbol{r}}_B^{(1)} &= \begin{bmatrix}630, & 150, & -5, & -5\end{bmatrix}^{\mathrm{T}}, \\
\hat{\boldsymbol{r}}_B^{(2)} &= \begin{bmatrix}700, & 100, & -5, & -5\end{bmatrix}^{\mathrm{T}}, \\
\hat{\boldsymbol{r}}_B^{(3)} &= \begin{bmatrix}250, & -350, & 5, & 5\end{bmatrix}^{\mathrm{T}}, \\
\hat{\boldsymbol{r}}_B^{(4)} &= \begin{bmatrix}250, & -430, & 5, & 5\end{bmatrix}^{\mathrm{T}},
\end{aligned}$$

with the same covariance $\mathbf{C}_B^{\boldsymbol{r}} = \mathrm{diag}\begin{bmatrix}.2, & .2, & .2, & .2\end{bmatrix}$ for all components. The birth components are assumed to have the same shape variable $\hat{p}_B = \begin{bmatrix}0, & 30, & 20\end{bmatrix}^{\mathrm{T}}$ with $\mathbf{C}_B^{\boldsymbol{p}} = \mathrm{diag}\begin{bmatrix}0.01, & 1, & 1\end{bmatrix}$. The measurement noise covariance is the same for all objects over the whole tracking interval: $\mathbf{C}_k^{\boldsymbol{v}} = \mathrm{diag}\begin{bmatrix}4, & 4\end{bmatrix}$. The initial existence probability of each birth component is 0.1. Poisson rates for birth components and existing objects are 15. The survival probability for an existing track is 0.99. The detection probability is 0.8. A constant velocity model is used for all objects. The process noise covariance matrices for all objects are the same, i.e., $\mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{r}}} = \mathrm{diag}\begin{bmatrix}10, & 10, & 10, & 10\end{bmatrix}$, and $\mathbf{C}_k^{\boldsymbol{w}^{\boldsymbol{p}}} = \mathrm{diag}\begin{bmatrix}0.01, & 0.1, & 0.1\end{bmatrix}$.

Result

Figure 4.14 depicts the tracking result of every third scan. It can be seen that the tracked trajectories are consistent with the ground truth and have no ID switches. The algorithm recognizes the newborn objects as soon as they entered the observation area. Our previous work [Yang2] gave similar results, but a distance-based clustering method was performed and had more time complexity. Furthermore, the algorithm in [Yang2] requires an additional module to maintain labels and handle occlusions.

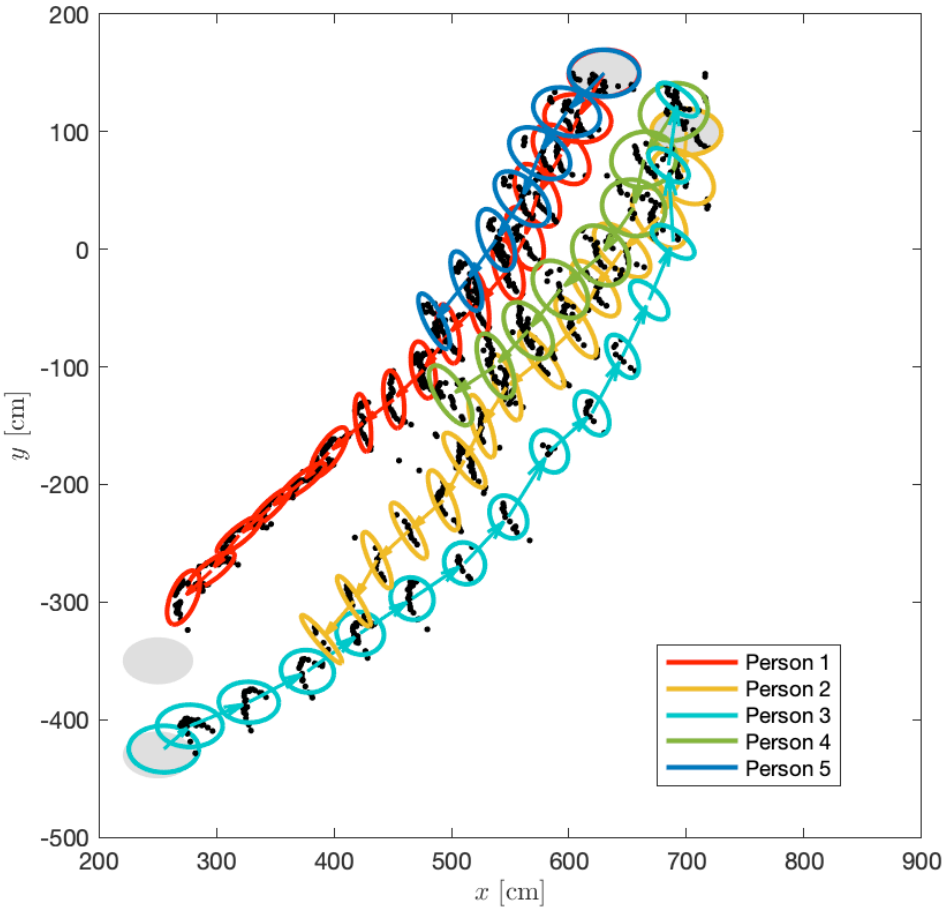


Figure 4.14: The pedestrian tracking result of MEM-JIPDA using the lidar data from the KITTI dataset. The filled gray ellipses are birth components. [Yang3] ©2020IEEE

4.9 Conclusion

In this chapter, we propose a MEOT tracker called MEM-JIPDA. MEM indicates the measurement model MEM-EKF* from Chapter 3. In addition, the key leads to the MEM-JIPDA is the data association method, which is inspired by the JIPDA filter proposed by Mušicki and Evas in [80]. This data association method calculates the marginal association probability for each measurement and existence probability for each object. The other modules, such as birth model, measurement update and track management, are designed to adapt this data association method for MEOT.

The proposed data association method

- requires no clustering or measurements enumeration,
- gives exact closed-form marginal association probabilities and existence probabilities by summing up the posterior probabilities of all measurement oriented joint associations,
- labels trajectories, which is inherited from JPDA and JIPDA, and
- employs the zero-inflated PPP model for modeling measurements.

The following factors are crucial for developing the closed-form marginal association probabilities:

1. *the incorporation of PPP into JPDA/JIPDA for extended objects.* The employment of PPP for EOT is widely accepted. The previous MEOT data association methods evaluate the likelihood of a measurement-set by taking the product of the predicted measurement likelihoods and the probability of generating the given numbers of measurements. The combination of JPDA and PPP facilitates independence among single measurement mappings. In the end, the likelihood of a measurement-set is proportional to the product of predicted measurement likelihoods and the respective Poisson rates.
2. *the design of the measurement-oriented joint association hypothesis.* The constraints of the joint association hypothesis of extended objects are: i) a measurement has one origin, ii) an object can be associated with multiple measurements. Therefore, a measurement-oriented joint association variable fulfills both constraints by the nature of its structure, and no additional conditions need to be checked.
3. *theorem 4.1, which avoids the enumeration of all joint assignments by converting the sum of a set of joint associations into a product of respective summation.* Using this theorem, the evaluation of all compatible joint association events in (4.22) and (4.25) requires no explicitly exhaustive enumeration of all joint association variables.

A sequential PDA update is adapted to combine the marginal association probabilities and the MEM-EKF* measurement update. Note that this sequential PDA measurement update is inherited from MEM-EKF*. Other measurement models, such as the random matrix [10, 5], are also possible.

Based on the data association method, a birth model is designed. The birth model assumes

that birth components are given in fixed locations with certain uncertainties. Compared with spontaneous birth modeling, this birth mechanism requires some background information on the surveillance area. More sophisticated birth models, such as the adaptive birth model [108], will be considered in the future.

5 Conclusion

Extended object tracking, which estimates the object extension in addition to kinematics based on a varying number of measurements, is a challenging research topic with increasing interest in many areas. This thesis is devoted to elliptical extended object tracking. The extension of an extended object is represented using an ellipse, and the measurements generated from an object are assumed to be randomly lying on the surface of the object. Based on these assumptions, three aspects of extended object tracking are addressed in this thesis:

- Metric for EOT comparison.

  For point object tracking, the Euclidean distance is the natural metric for evaluation. As the extension estimation is one of the crucial tasks in EOT, there is no consistently used metric for assessing extended object trackers yet. Chapter 2 explains the problem of evaluating EOT and investigated the measures that can be used for assessing elliptical shapes. Based on the discussion and insight that are gained from simulations, the Gaussian Wasserstein distance is proposed for assessing the performance of elliptical extended object trackers.

- Extension estimation.

  Chapter 3 formulates an explicit measurement model with orientation and semi-axes lengths and presents a single extended object tracker that estimates the object extension using a tailored Kalman filter. Compared with a matrix representation for the elliptical extension, the parameterization with orientation and axes lengths facilitates the dynamic modeling of individual shape parameters in many applications. For example, a coordinate turn model, which elaborates the dependency between extension transformation and kinematic state explicitly, can be easily constructed. Furthermore, a full joint covariance matrix for the estimate of orientation and axes lengths is maintained. A tailored extended Kalman filter is derived and results in compact update formulas.

- Multiple extended object tracking.

  One of the bottlenecks of MOT for point objects is the data association problem between measurements and objects. As one extended object can generate a varying number of measurements, the data association problem in EOT is even computational challenging. The tracker proposed in Chapter 3 assumes the origins of the measurements are known. In Chapter 4, this assumption is relaxed. Besides, the number of tracked objects is varying and unknown. The state-of-the-art methods bypass the data association problem for extended objects using clustering or measurement enumeration, which has high computational complexity. Chapter 4 proposes a new data association method without enumerating measurement partitions or clustering measurements. Based on the proposed data association method, an efficient MEOT is introduced by combining the shape tracker proposed in Chapter 3. The effectiveness of the new MEOT tracker is demonstrated using simulation and real lidar data.

One direction of the future work will be investigating the explicit measurement model with the non-elliptical shapes, such as rectangles or arbitrary shapes. As the data association method proposed in the multiple extended object tracker is presented using marginal association probability and requires no explicit measurement partitioning, one possible direction of the future research is an investigation on combining the data association method with a learned measurement model using neural networks. Furthermore, exploring the fusion of the extended object estimates from multiple and heterogeneous sensors, such as camera, radar, and lidar, has enormous practical and research potential.

Acronyms

CPHD	Cardinalized Probability Hypothesis Density	64
EOT	Extended Object Tracking	3
EM	Expectation Maximization	66
GPDA	Generalized Probabilistic Data Association	65
GJPDA	Generalized Joint Probabilistic Data Association	65
GNN	Global Nearest Neighbour	64
GGIW-PHD	Gamma Gaussian Wishart Probability Hypothesis Density	92
LMB	Labeled Multi-Bernoulli	64
KF	Kalman Filter	8
IPDA	Integrated Probabilistic Data Association	64
IoU	Intersection over Union	16
JPDA	Joint Probabilistic Data Association	63
JIPDA	Joint Integrated Probabilistic Data Association	62
MEM	Multiplicative Error Model	37
MEM-EKF*	Multiplicative Error Model-Extended Kalman Filtering*	29
MHT	Multi-Hypothesis Tracking	64
MD-PDA	Multiple-Detection Probabilistic Data Association	65
MD-JPDA	Multiple-Detection Joint Probabilistic Data Association	65
MOT	Multiple Object Tracking	7
MEOT	Multiple Extended Object Tracking	7
NN	Nearest Neighbour	64
OSPA	Optimal Sub-Pattern Assignment	9
PDA	Probabilistic Data Association	64
PHD	Probability Hypothesis Density	62

PPP Poisson Point Process 6
PMBM Poisson multi-Bernoulli mixture 64
PMHT Probabilistic Multi-Hypothesis Tracking 64
RFS Random Finite sets 64
RHM Random Hypersurface Model 9
RMSE Root Mean Squared Error 12
SPD Symmetric Positive Definite 8

Nomenclature

$\otimes$	Kronecker product
$\mathbf{A}$	a matrix or a set
$\boldsymbol{a}$	a random vector
$\hat{\boldsymbol{a}}$	an estimate of $\boldsymbol{a}$
$\hat{\boldsymbol{a}}^{(i)}$	the i-th estimate of $\boldsymbol{a}$
$\mathbf{C}^{\boldsymbol{h}}$	covariance of multiplicative noise
$\text{cov}\{\boldsymbol{x}, \boldsymbol{y}\}$	covariance between $\boldsymbol{x}$ and $\boldsymbol{y}$
$\mathbf{C}_k^{\boldsymbol{w}^r}$	process noise covariance of variable $\boldsymbol{r}$ at time step k
$\mathbf{C}_k^{\boldsymbol{x}(\tau,m)}$	m-th updated covariance of $\boldsymbol{x}$ for object τ at time step k
$\mathbf{C}_k^{\boldsymbol{x}+}$	updated covariance of $\boldsymbol{x}$ at time step k
$\mathbf{C}_k^{\boldsymbol{x}}$	covariance of $\boldsymbol{x}$ at time step k
$\mathbf{C}_k^{\boldsymbol{xy}(i)}$	i-th updated cross covariance of $\boldsymbol{x}$ and $\boldsymbol{y}$ at time step k
$\mathbf{C}_k^{\boldsymbol{xy}}$	cross covariance of $\boldsymbol{x}$ and $\boldsymbol{y}$ at time step k
$d(\boldsymbol{x}_1, \boldsymbol{x}_2)$	distance between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$
$\text{diag}[\cdot]$	diagonal matrix

$\mathrm{E}(\boldsymbol{x})$ expectation of $\boldsymbol{x}$

$\mathbf{H}_k$ measurement matrix at time step k

$\boldsymbol{h}_k^{(i)}$ i-th multiplicative noise at time step k

$\mathbf{I}_n$ n dimensional identity matrix

$\mathbf{J}_f$ Jacobian matrix of f

$l_{k,1}, l_{k,2}$ semi-axes lengths of an elliptical object at time step k

$l_{m\tau}$ predicted likelihood of m-th measurement given object τ

$\mathbf{M}_k$ set of measurements indices at time step k

M_k number of measurements at time step k

$\boldsymbol{m}_k, \dot{\boldsymbol{m}}_k$ location and velocity of an extended object

$\mathbf{N}_k$ set of object indices at times step k

$\mathbf{N}_{B,k}$ set of birth component indices at times step k

$\mathbf{N}_{S,k}$ set of survived object indices at times step k

N_k number of objects at times step k

$\mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$ Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$

$\mathrm{Pois}(\cdot \mid \lambda)$ Poisson probability mass function with rate λ

$\mathcal{P}_{\boldsymbol{x}}$ paraterization of an ellipse with center, orientation and semi-axes lengths

$\mathbb{P}(\mathbf{N})$ powerset of $\mathbf{N}$

$\mathbb{P}(\mathbf{N})^{-}$ powerset of $\mathbf{N}$ excluding the set $\mathbf{N}$

P_D detection probability

$P_{D,k}^{(\tau)}$ detection probability of object τ at time step k

P_G gating probability

P_S survival probability

$P_{\exists,k}^{(\tau)}$ existence probability of object τ at time step k

$P_{\exists,k}^{(\tau)+}$ updated existence probability of object τ at time step k

$P_{\nexists,k}^{(\tau)}$ non-existence probability of object τ at time step k

$P_{\nexists,k}^{(\tau)+}$ updated non-existence probability of object τ at time step k

$p(x)$ probability density function of x

$P(x)$ probability of x

$\boldsymbol{p}_k$ shape variable of an elliptical extended object at time step k

$\boldsymbol{r}_k$ kinematic state of an extended object at time step k

$\hat{\boldsymbol{r}}_k^{(m)}$ m-th update of kinematic state of an object at time step k

$\hat{\boldsymbol{r}}_k^{(m,\tau)}$ m-th update of kinematic state of object τ at time step k

$\hat{\boldsymbol{p}}_k^{(m)}$ m-th update of shape variable of an object at time step k

$\hat{\boldsymbol{p}}_k^{(m,\tau)}$ m-th update of shape variable state of object τ at time step k

$\mathbf{S}_k$ Square root matrix of an elliptical shape matrix at time step k

$\mathbf{T}_{\boldsymbol{\theta}_k}$ numbers of measurements assigned to each object by $\boldsymbol{\theta}_k$

$\mathbf{U}_k$ set of undetected and non-existent object indices at time step k

$\boldsymbol{v}_k$ measurement noise at time step k

$\text{vect}\{\mathbf{A}\}$ converts a matrix into a column vector

$\boldsymbol{w}_k^{\boldsymbol{r}}$ process noise of variable $\boldsymbol{r}$

$\boldsymbol{w}_k$ process noise at time step k

$\mathbf{X}_k$ set of objects at time step k

$\hat{\mathbf{X}}_k$ set of estimated objects at time step k

$\boldsymbol{x}_k$ object state at time step k

$\boldsymbol{x}_{1:k}$ object states from time step 1 to k

$\hat{\boldsymbol{x}}_k^{+}$ mean of the updated state at times step k

$\hat{\boldsymbol{x}}_k$ mean of the predicted state at time step k

$\mathbf{Y}_k$	set of measurements at time step k
$\mathcal{Y}_k^{(\tau)}$	set of measurements assigned to object τ
$\boldsymbol{y}_k$	measurement at time step k
$\boldsymbol{y}_{1:k}$	measurements from time step 1 to k
$\boldsymbol{y}_k^{(i)}$	i-th measurement at time step k
$\bar{\boldsymbol{y}}_k^{(\tau,m)}$	expected m-th measurement of object τ at time step k
$\bar{\boldsymbol{y}}_k^{(i)}$	expected i-th measurement at time step k
$\boldsymbol{z}_k^{(i)}$	the measurement source of $\boldsymbol{y}_k^{(i)}$
α_k	orientation of an ellipse at time step k
$\beta_{m\tau}$	marginal association probability of m-th measurement being assigned to object τ given the existence of object τ
δ_{initiate}	threshold for the initiation of a track
$\delta_{\text{termiante}}$	threshold for the termination of a track
$\boldsymbol{\Theta}_{\mathbf{N}}^{\mathbf{M}}$	set of joint association variables from $\mathbf{M}$ to $\mathbf{N}$
$\boldsymbol{\theta}_k$	joint association variable
θ_m	object index that m-th measurement is associated with
$\lambda_k^{(\tau)}$	Poisson rate of object τ at time step k
$\boldsymbol{\xi}_k^{(i)}$	i-th pseudo-measurement
$\bar{\boldsymbol{\xi}}_k^{(i)}$	expected i-th pseudo-measurement
$\bar{\boldsymbol{\xi}}_k^{(\tau,m)}$	expected m-th pseudo-measurement of object τ
ρ	spatial density of clutter
$\boldsymbol{\Sigma}_{\boldsymbol{x}_k}$	shape matrix of $\boldsymbol{x}_k$

Own Publications

Journal Publications

[Yang1] **Shishan Yang** and Marcus Baum. "Tracking the Orientation and Axes Lengths of an Elliptical Extended Object". In: *IEEE Transactions on Signal Processing* 67.18 (2019), pp. 4720–4729. ISSN: 1053-587X. DOI: 10.1109/TSP.2019.2929462.

[Yang2] **Shishan Yang**, Florian Teich, and Marcus Baum. "Network Flow Labeling for Extended Target Tracking PHD filters". In: *IEEE Transactions on Industrial Informatics* 15.7 (2019), pp. 4164–4171. ISSN: 1551-3203. DOI: 10.1109/TII.2019.2898992.

Conference Publications

[Yang3] **Shishan Yang**, Laura M. Wolf, and Marcus Baum. "Marginal Association Probabilities for Multiple Extended Objects without Enumeration of Measurement Partitions ". In: *Proceedings of the 23rd International Conference on Information Fusion (Fusion 2020)*. Virtual, July 2020. DOI: 10.23919/FUSION45008.2020.9190500.

[Yang4] Kolja Thormann, **Shishan Yang**, and Marcus Baum. "A Comparison of Kalman Filter-based Approaches for Elliptic Extended Object Tracking". In: *Proceedings of the 23rd International Conference on Information Fusion (Fusion 2020)*. Virtual, July 2020. DOI: 10.23919/FUSION45008.2020.9190375.

[Yang5] **Shishan Yang**, Kolja Thormann, and Marcus Baum. "Linear-Time Joint Probabilistic Data Association for Multiple Extended Object Tracking". In: *2018 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2018)*. Sheffield, United Kingdom, July 2018. DOI: 10.1109/SAM.2018.8448430.

[Yang6] Thorben Janz, Andreas Leich, Marek Junghans, Kay Gimm, **Shishan Yang**, and Marcus Baum. "Post-Processing of Multi-Target Trajectories for Traffic Safety Analysis". In: *Proceedings of the 21st International Conference on Information Fusion (FUSION 2018)*. Cambridge, United Kingdom, July 2018. DOI: 10.23919/ICIF.2018.8455280.

[Yang7] **Shishan Yang** and Marcus Baum. "Extended Kalman Filter for Extended Object Tracking". In: *Proceedings of the 42nd IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2017)*. New Orleans, USA, Mar. 2017. DOI: 10.1109/ICASSP.2017.7952985.

[Yang8] Florian Teich, **Shishan Yang**, and Marcus Baum. "GM-PHD filter for Multiple Extended Object Tracking based on the Multiplicative Error Shape Model and Network Flow Labeling". In: *Proceedings of the IEEE Intelligent Vehicles Symposium (IV 2017)*. Redondo Beach, CA, USA, June 2017. DOI: 10.1109/IVS.2017.7995691.

[Yang9] **Shishan Yang**, Marcus Baum, and Karl 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. DOI: 10.1109/MFI.2016.7849541.

[Yang10] **Shishan Yang** and Marcus Baum. "Second-Order Extended Kalman Filter for Extended Object and Group Tracking". In: *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*. Heidelberg, Germany, July 2016.

[Yang11] Marcus Baum, **Shishan Yang**, and Uwe D Hanebeck. "The Kernel-SME Filter with False and Missing Measurements". In: *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*. Heidelberg, Germany, July 2016.

[Yang12] Stefan Breuers, **Shishan Yang**, Markus Mathias, and Bastian Leibe. "Exploring Bounding Box Context for Multi-Object Tracker Fusion". In: *IEEE Winter Conference on Applications of Computer Vision (WACV 2016)*. New York, USA, 2016. DOI: 10.1109/WACV.2016.7477564.

Bibliography

[1] Holger Caesar, Varun Bankiti, Alex H. Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. "NuScenes: A Multimodal Dataset for Autonomous Driving". In: *arXiv preprint arXiv:1903.11027* (2019).

[2] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. "Vision Meets Robotics: The KITTI Dataset". In: *The International Journal of Robotics Research* 32.11 (2013), pp. 1231–1237.

[3] Marcus Baum. *Simultaneous Tracking and Shape Estimation of Extended Objects*. KIT Scientific Publishing, 2013.

[4] Karl Granström, Marcus Baum, and Stephan Reuter. "Extended Object Tracking: Introduction, Overview and Applications". In: *ISIF Journal of Advances in Information Fusion* 12.2 (Dec. 2017). ISSN: 1557-6418.

[5] Michael Feldmann, Dietrich Fränken, and Wolfgang Koch. "Tracking of Extended Objects and Group Targets using Random Matrices". In: *IEEE Transactions on Signal Processing* 59.4 (2011), pp. 1409–1420. ISSN: 1053-587X. DOI: 10.1109/TSP.2010.2101064.

[6] Marcus Baum, Benjamin Noack, and Uwe D. Hanebeck. "Extended Object and Group Tracking with Elliptic Random Hypersurface Models". In: *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*. Edinburgh, United Kingdom, July 2010. DOI: 10.1109/ICIF.2010.5711854.

[7] Donka Angelova, Lyudmila Mihaylova, Nikolay Petrov, and Amadou Gning. "A Convolution Particle Filtering Approach for Tracking Elliptical Extended Objects". In: *Proceedings of the 16th International Conference on Information Fusion (Fusion 2013)*. IEEE. 2013, pp. 1542–1549.

[8] Süleyman F. Kara and Emre Özkan. "Multi-Ellipsoidal Extended Target Tracking Using Sequential Monte Carlo". In: *2018 21st International Conference on Information Fusion (FUSION)*. 2018, pp. 1–8. DOI: 10.23919/ICIF.2018.8455436.

[9] Kolja Thormann and Marcus 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.

[10] Wolfgang Koch. "Bayesian Approach to Extended Object and Cluster Tracking using Random Matrices". In: *IEEE Transactions on Aerospace and Electronic Systems* 44.3 (July 2008), pp. 1042–1059. ISSN: 0018-9251. DOI: 10.1109/TAES.2008.4655362.

[11] Marcus Baum and Uwe D. Hanebeck. "Extended Object Tracking with Random Hypersurface Models". In: *IEEE Transactions on Aerospace and Electronic Systems* 50 (Jan. 2014), pp. 149–159. DOI: 10.1109/TAES.2013.120107.

[12] Karl Granström, Christian Lundquist, and Umut 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, 2011.

[13] Hauke Kaulbersch, Jens Honer, and Marcus Baum. "A Cartesian B-Spline Vehicle Model for Extended Object Tracking". In: *21st International Conference on Information Fusion (FUSION 2018)*. Cambridge, United Kingdom, July 2018.

[14] Ted J. Broida and Rama Chellappa. "Estimating the Kinematics and Structure of a Rigid Object from a Sequence of Monocular Images". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13.6 (1991), pp. 497–513. ISSN: 0162-8828. DOI: 10.1109/34.87338.

[15] Jean C. Dezert. "Tracking Maneuvering and Bending Extended Target in Cluttered Environment". In: *Signal and Data Processing of Small Targets*. Vol. 3373. Orlando, USA, 1998, pp. 283–294.

[16] Lars Hammarstrand, Lennart Svensson, Fredrik Sandblom, and Joakim Sörstedt. "Extended Object Tracking using a Radar Resolution Model". In: *IEEE Transactions on Aerospace and Electronic Systems* 48.3 (2012), pp. 2371–2386.

[17] Ronald P. S. Mahler. *Advances in Statistical Multisource-Multitarget Information Fusion*. Artech House, 2014.

[18] Kevin Gilholm, Simon Godsill, Simon Maskell, and Dvaivd Salmond. "Poisson Models for Extended Target and Group Tracking". In: *SPIE: Signal and Data Processing of Small Targets*. 2005.

[19] Kevin Gilholm and David Salmond. "Spatial Distribution Model for Tracking Extended Objects". In: *IEE Proceedings on Radar, Sonar and Navigation* 152.5 (Oct. 2005), pp. 364–371. ISSN: 1350-2395.

[20] Karl Granström, Antonio Natale, Paolo Braca, Giovanni Ludeno, and Francesco Serafino. "Gamma Gaussian Inverse Wishart Probability Hypothesis Density for Extended Target Tracking using X-Band Marine Radar Data". In: *IEEE Transactions on Geoscience and Remote Sensing* 53.12 (Dec. 2015), pp. 6617–6631.

[21] Michael Beard, Stephan Reuter, Karl Granström, Ba-Tuong Vo, Ba-Ngu Vo, and Alexander Scheel. "Multiple Extended Target Tracking with Labelled Random Finite Sets". In: *IEEE Transactions on Signal Processing* 64.7 (Apr. 2016), pp. 1638–1653.

[22] Branko Ristic and Jamie Sherrah. "Bernoulli Ffilte for Joint Detection and Tracking of an Extended object in clutter". In: *IET Radar, Sonar Navigation* (2012).

[23] Branko Ristic, Ba-Tuong Vo, Ba-Ngu Vo, and Alfonso Farina. "A Tutorial on Bernoulli Filters: Theory, Implementation and Applications". In: *IEEE Transactions on Signal Processing* 61.13 (2013), pp. 3406–3430.

[24] Anna Petrovskaya and Sebastian Thrun. "Model Based Vehicle Tracking for Autonomous Driving in Urban Environments". In: *Auton Robot* 26 (2009), pp. 123–139.

[25] Gemine Vivone and Paolo Braca. "Joint Probabilistic Data Association Tracker for Extended Target Tracking Applied to X-Band Marine Radar Data". In: *IEEE Journal of Oceanic Engineering* 41.4 (Oct. 2016), pp. 1007–1019.

[26] Marcus Baum and Uwe D. Hanebeck. "Shape Tracking of Extended Objects and Group Targets with Star-Convex RHMs". In: *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*. Chicago, Illinois, USA, July 2011.

[27] Niklas Wahlström and Emre Özkan. "Extended Target Tracking Using Gaussian Processes". In: *IEEE Transactions on Signal Processing* 63.16 (2015), pp. 4165–4178.

[28] Jian Lan and X. Rong Li. "Tracking of Extended Object or Target Group Using Random Matrix: New Model and Approach". In: *IEEE Transactions on Aerospace and Electrical Systems* 52.6 (2016), pp. 2973–2988.

[29] Felix Govaers. "On Independent Axes Estimation for Extended Target Tracking". In: *2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*. 2019, pp. 1–6.

[30] Dominic Schuhmacher, Ba-Tuong Vo, and Ba-Ngu Vo. "A Consistent Metric for Performance Evaluation of Multi-Object Filters". In: *IEEE Transactions on Signal Processing* 56.8 (Aug. 2008), pp. 3447 –3457. ISSN: 1053-587X. DOI: 10.1109/TSP.2008.920469.

[31] Yaakov Bar-Shalom, Fred Daum, and Jim Huang. "The Probabilistic Data Association Filter". In: *IEEE Control Systems Magazine* 29.6 (2009), pp. 82–100. ISSN: 0272-1708. DOI: 10.1109/MCS.2009.934469.

[32] Ba-Ngu Vo, Ba-Tuong Vo, and Hung G. Hoang. "An Efficient Implementation of the Generalized Labeled Multi-Bernoulli Filter". In: *IEEE Transactions on Signal Processing* 65.8 (2017), pp. 1975–1987. ISSN: 1941-0476. DOI: 10.1109/TSP.2016.2641392.

[33] Jian Lan and X. Rong Li. "Extended-Object or Group-Target Tracking Using Random Matrix With Nonlinear Measurements". In: *IEEE Transactions on Signal Processing* 67.19 (2019), pp. 5130–5142.

[34] Nathan James Bartlett, Christopher Renton, and Adrian G. Wills. "A Closed-Form Prediction Update for Extended Target Tracking Using Random Matrices". In: *IEEE Transactions on Signal Processing* 68 (2020), pp. 2404–2418.

[35] Karl Granström, Maryam Fatemi, and Lennart Svensson. "Poisson Multi-Bernoulli Mixture Conjugate Prior for Multiple Extended Target Filtering". In: *IEEE Transactions on Aerospace and Electronic Systems* 56.1 (2020), pp. 208–225.

[36] Karl Granström and Jakob Bramstång. "Bayesian Smoothing for the Extended Object Random Matrix Model". In: *IEEE Transactions on Signal Processing* 67.14 (2019), pp. 3732–3742.

[37] Tim Baur, Julian Böhler, Stefan Wirtensohn, and Johannes Reuter. "Tracking of Spline Modeled Extended Targets Using a Gaussian Mixture PHD Filter". In: *2019 22th International Conference on Information Fusion (FUSION)*. 2019, pp. 1–8.

[38] Keni Bernardin and Rainer Stiefelhagen. "Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics". In: *EURASIP Journal on Image and Video Processing* 2008 (2008), pp. 1–10.

[39] John R. Hoffman and Ronald P. S. Mahler. "Multitarget Miss Distance via Optimal Assignment". In: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 34.3 (2004), pp. 327–336.

[40] Harold W. Kuhn. "The Hungarian Method for the Assignment Problem". In: *Naval Research Logistics Quarterly* 2 (1955), pp. 83–97.

[41] Branko Ristic, Ba-Ngu Vo, Daniel Clark, and Ba-Tuong Vo. "A Metric for Performance Evaluation of Multi-Target Tracking Algorithms". In: *IEEE Transactions on Signal Processing* 59.7 (2011), pp. 3452–3457.

[42] Ángel F. García-Fernández, Abu Sajana Rahmathullah, and Lennart Svensson. "A Metric on the Space of Finite Sets of Trajectories for Evaluation of Multi-Target Tracking Algorithms". In: *IEEE Transactions on Signal Processing* 68 (2020), pp. 3917–3928.

[43] Marcus Baum, Florian Faion, and Uwe D. Hanebeck. "Modeling the Target Extent with Multiplicative Noise". In: *Proceedings of the 15th International Conference on Information Fusion (Fusion 2012)*. Singapore, July 2012.

[44] Jian Lan and X. Rong Li. "Extended Object or Group Target Tracking Using Random Matrix with Nonlinear Measurements". In: *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*. Heidelberg, Germany, 2016.

[45] Christian Lundquist, Karl Granström, and Umut Orguner. "An Extended Target CPHD Filter and a Gamma Gaussian Inverse Wishart Implementation". In: *IEEE Journal of Selected Topics in Signal Processing* 7.3 (2013), pp. 472–483.

[46] Karl Granström and Umut Orguner. "A New Prediction for Extended Targets with Random Matrices". In: *IEEE Transactions on Aerospace and Electrical Systems* 50.2 (2014).

[47] Ian L. Dryden, Alexey Koloydenko, and Diwei Zhou. "Non-Euclidean Statistics for Covariance Matrices, with Applications to Diffusion Tensor Imaging," in: *The Annals of Applied Statistics* 3.3 (2009), pp. 1102–1123.

[48] Gabriela Csurka, Diane Larlus, and Florent Perronnin. "What is a Good Evaluaton Measure for Semantic Segmentation?" In: *24th British Machine Vision Conference (BMVC 2013)*. Bristol, 2013.

[49] Mark Everingham, S. M. Ali Eslami, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman. "The Pascal Visual Object Classes Challenge: A Retrospective". In: *International Journal of Computer Vision* 111.1 (2015), pp. 98–136.

[50] Jacinto Nascimento and Jorge Marques. "Performance Evaluation of Object Detection Algorithms for Video Surveillance". In: *IEEE Transactions on Multimedia* 8.4 (2006), pp. 761–773.

[51] C. Lawrence Zitnick and Piotr Dollár. "Edge Boxes: Locating Object Proposals from Edges". In: *The 10th European Conference on Computer Vision (ECCV 2008)*. Marseille, France, 2008.

[52] Anton Milan, Konrad Schindler, and Stefan Roth. "Challenges of Ground Truth Evaluation of Multi-Target Tracking". In: *Proc. of the CVPR 2013 Workshop on Ground Truth - What is a good dataset?* Portland, Oregon, USA, 2013.

[53] Anton Milan, Konrad Schindler, and Stefan Roth. "Multi-Target Tracking by Discrete-Continuous Energy Minimization". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38.10 (2016), pp. 2054–2068.

[54] Solomon Kullback and Richard A. Leibler. "On Information and Sufficiency". In: *Annals of Mathematical Statistics* 22.1 (1951), pp. 79–86.

[55] Julian Böhler, Tim Baur, Stefan Wirtensohn, and Johannes Reuter. "Stochastic Partitioning for Extended Object Probability Hypothesis Density Filters". In: *2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*. 2019, pp. 1–6.

[56] John R. Hoffman and Ronald P.S. Mahler. "Multitarget Miss Distance and its Applications". In: *Proceedings of the 5th International Conference on Information Fusion*. Vol. 1. July 2002, 149–155 vol.1.

[57] Cédric Villani. *Optimal Transport*. Springer Verlag, 2009.

[58] Clark R. Givens and Rae Michael Shortt. "A Class of Wasserstein Metrics for Probability Distributions". In: *The Michigan Mathematical Journal* 31.2 (1984), pp. 231–240.

[59] K. Thormann and M. Baum. "Fusion of Elliptical Extended Object Estimates Parameterized with Orientation and Axes Lengths". In: *IEEE Transactions on Aerospace and Electronic Systems* (2021), pp. 1–1. DOI: 10.1109/TAES.2021.3057651.

[60] Remco Veltkamp and Michiel Hagedoorn. "Shape Similarity Measures, Properties and Constructions". In: *Advances in Visual Information Systems*. Ed. by Robert Laurini. Vol. 1929. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2000, pp. 133–153.

[61] Hauke Kaulbersch, Marcus Baum, and Peter Willett. "EM Approach for Tracking Star-Convex Extended Objects". In: *Proceedings of the 20th International Conference on Information Fusion (Fusion 2017)*. Xi'an, P.R. China, July 2017. DOI: 10.23919/ICIF.2017.8009884.

[62] Jaco Vermaak, Norikazu Ikoma, and Simon Godsill. "Sequential Monte Carlo Framework for Extended Object Tracking". In: *IEE Proceedings on Radar, Sonar and Navigation* 152.5 (2005), pp. 353–363. ISSN: 1350-2395. DOI: 10.1049/ip-rsn:20045044.

[63] Jian Lan and X. Rong Li. "Tracking of Extended Object or Target Group Using Random Matrix - Part I: New Model and Approach". In: *Proceedings of the 15th International Conference on Information Fusion (Fusion 2012)*. July 2012, pp. 2177–2184.

[64] Yvo Boers, Hans Driessen, Johan Torstensson, Mikael Trieb, Rickard Karlsson, and Fredrik Gustafsson. "Track-Before-Detect Algorithm for Tracking Extended Targets". In: *IEE Proceedings on Radar, Sonar and Navigation* 153.4 (Aug. 2006), pp. 345–351. ISSN: 1350-2395.

[65] Karl Granström and Christian Lundquist. "On the Use of Multiple Measurement Models for Extended Target Tracking". In: *16th International Conference on Information Fusion (Fusion 2013)*. 2013.

[66] Karl Granström, Stefan Reuter, Daniel Meissner, and Alexander Scheel. "A Multiple Model PHD Approach to Tracking of Cars under an Assumed Rectangular Shape". In: *Proceedings of the 17th International Conference on Information Fusion (Fusion 2014)*. Salamanca, Spain, July 2014.

[67] Christina Knill, Alexander Scheel, and Klaus Dietmayer. "A Direct Scattering Model for Tracking Vehicles with High-Resolution Radars". In: *2016 IEEE Intelligent Vehicles Symposium (IV)*. Gothenburg, Sweden, 2016.

[68] Marcus Baum, Vesa Klumpp, and Uwe D. Hanebeck. "A Novel Bayesian Method for Fitting a Circle to Noisy Points". In: *Proceedings of the 13th International Conference on Information Fusion (Fusion 2010)*. Edinburgh, United Kingdom, July 2010. DOI: 10.1109/ICIF.2010.5711884.

[69] Nikolay Petrov, Lyudmila Mihaylova, Amadou Gning, and Donka Angelova. "A Novel Sequential Monte Carlo Approach for Extended Object Tracking based on Border Parameterisation". In: *Proceedings of the 14th International Conference on Information Fusion (Fusion 2011)*. Chicago, Illinois, USA, July 2011.

[70] Nikolay Petrov, Amadou Gning, Lyudmila Mihaylova, and Donka Angelova. "Box Particle Filtering for Extended Object Tracking". In: *Proceedings of the 15th International Conference on Information Fusion (Fusion 2012)*. Singapore: IEEE, July 2012, pp. 82–89. ISBN: 978-1-4673-0417-7.

[71] Marcus Baum, Michael Feldmann, Dietrich Fränken, Uwe D. Hanebeck, and Wolfgang Koch. "Extended Object and Group Tracking: A Comparison of Random Matrices and Random Hypersurface Models". In: *Proceedings of the IEEE ISIF Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF 2010)*. Leipzig, Germany, Oct. 2010.

[72] Michael Kaess, Ananth Ranganathan, and Frank Dellaert. "iSAM: Incremental Smoothing and Mapping". In: *IEEE Transactions on Robotics* 24.6 (Dec. 2008), pp. 1365–1378.

[73] Gemine Vivone, Paolo Braca, Karl Granström, Antonio Natale, and Jocelyn Chanussot. "Converted Measurements Random Matrix Approach to Extended Target Tracking Using X-band Marine Radar Data". In: *18th International Conference on Information Fusion (Fusion 2015)*. Washington, DC, USA, July 2015, pp. 976–983.

[74] Umut Orguner. "A Variational Measurement Update for Extended Target Tracking with Random Matrices". In: *IEEE Transactions on Signal Processing* 60.7 (2012), pp. 3827 –3834. ISSN: 1053-587X. DOI: 10.1109/TSP.2012.2192927.

[75] Tohid Ardeshiri, Umut Orguner, and Fredrik Gustafsson. "Bayesian Inference via Approximation of Log-likelihood for Priors in Exponential Family". In: *Arxiv*. Vol. CoRR, vol. abs/1510.01225. 2015.

[76] Lyudmila Mihaylova, Avishy Carmi, François Septier, Amadou Gning, Sze Pang, and Simon Godsill. "Overview of Bayesian Sequential Monte Carlo Methods for Group and Extended Object Tracking". In: *Digital Signal Processing* 25 (Feb. 2014), pp. 1–16.

[77] Jian Lan and Xiao R. Li. "Nonlinear Estimation by LMMSE-Based Estimation With Optimized Uncorrelated Augmentation". In: *IEEE Transactions on Signal Processing* 63.16 (2015), pp. 4270–4283. ISSN: 1053-587X. DOI: 10.1109/TSP.2015.2437834.

[78] Leon Isserlis. "On A Formula For The Product-Moment Coefficient of Any Order of a Normal Frequency Distribution in Any Number of Variables". In: *Biometrika* 12.1-2 (1918), pp. 134–139.

[79] Gian C. Wick. "The Evaluation of the Collision Matrix". In: *Physical Review* 80.2 (1950), pp. 268–272.

[80] Darko Mušicki and Robin Evans. "Joint Integrated Probabilistic Data Association - JIPDA". In: *Proceedings of the Fifth International Conference on Information Fusion. FUSION 2002. (IEEE Cat.No.02EX5997)*. Vol. 2. 2002, 1120–1125 vol.2. DOI: 10.1109/ICIF.2002.1020938.

[81] Samuel S. Blackman and Robert F. Popoli. *Design and Analysis of Modern Tracking Systems*. Arthech House, 1999.

[82] Dimitri Bertsekas. "The Auction Algorithm: A Distributed Relaxation Method for the Assignment Problem". In: *Annals of Operations Research* 14 (Dec. 1988), pp. 105–123. DOI: 10.1007/BF02186476.

[83] Katta G. Murty. "An Algorithm for Ranking all the Assignments in Order of Increasing Cost". In: *Operations Research* 16.3 (1968), pp. 682–687. ISSN: 0030364X, 15265463.

[84] Kuo-Chu Chang and Yaakov Bar-Shalom. "Joint Probabilistic Data Association for Multitarget Tracking with Possibly Unresolved Measurements and Maneuvers". In: *IEEE Transactions on Automatic Control* 29.7 (1984), pp. 585–594. ISSN: 2334-3303. DOI: 10.1109/TAC.1984.1103597.

[85] Darko Mušicki, Robin Evans, and Srdjan Stankovic. "Integrated Probabilistic Data Association". In: *IEEE transactions on Automatic Control* 39.6 (1994), pp. 1237–1241.

[86] Donald Reid. "An Algorithm for Tracking Multiple Targets". In: *IEEE Transactions on Automatic Control* 24.6 (1979), pp. 843–854.

[87] Ingemar J. Cox and Sunita L. Hingorani. "An Efficient Implementation and Evaluation of Reid's Multiple Hypothesis Tracking Algorithm for Visual Tracking". In: *Proceedings of 12th International Conference on Pattern Recognition*. Vol. 1. 1994, 437–442 vol.1.

[88] Ronald P. S. Mahler. "Multitarget Bayes Filtering via First-Order Multitarget Moments". In: *IEEE Transactions on Aerospace and Electronic Systems* 39.4 (Oct. 2003), pp. 1152 –1178. ISSN: 0018-9251. DOI: 10.1109/TAES.2003.1261119.

[89] Ronald P. S. Mahler. *Statistical Multisource-Multitarget Information Fusion*. Boston, Mass.: Artech House, 2007. ISBN: 978-1-59693-092-6 ; 1-59693-092-6.

[90] Ba-Tuong Vo, Ba-Ngu Vo, and A. Cantoni. "Analytic Implementations of the Cardinalized Probability Hypothesis Density Filters". In: *IEEE Transactions on Signal Processing* 55.7 (2007), pp. 3553 –3567. ISSN: 1053-587X. DOI: 10.1109/TSP.2007.894241.

[91] Karl Granström, Christian Lundquist, and Omut Orguner. "Extended Target Tracking Using a Gaussian-Mixture PHD Filter". In: *IEEE Transactions on Aerospace and Electronic Systems* 48.4 (2012), pp. 3268–3286.

[92] Karl Granström and Umut Orguner. "A PHD Filter for Tracking Multiple Extended Targets Using Random Matrices". In: *IEEE Transactions on Signal Processing* 60.11 (Nov. 2012), pp. 5657–5671. ISSN: 1053-587X. DOI: 10.1109/TSP.2012.2212888.

[93] Biruk Habtemariam, Ratnasigham Tharmarasa, Thayaparan Thayaparan, Mahendra Mallick, and Thia Kirubarajan. "A Multiple-Detection Joint Probabilistic Data Association Filter". In: *IEEE Journal of Selected Topics in Signal Processing* 7.3 (2013), pp. 461–471.

[94] Zhenzhen Su, Hongbing Ji, and Yongquan Zhang. "Loopy Belief Propagation based Data Association for Extended Target Tracking". In: *Chinese Journal of Aeronautics* (Mar. 2020). DOI: 10.1016/j.cja.2020.01.004.

[95] Ronald P. S. Mahler. "PHD Filters for Nonstandard Targets, I: Extended Targets". In: *Proceedings of the 12th International Conference on Information Fusion (Fusion 2009)*. Seattle, Washington, July 2009.

[96] Roy Streit. "JPDA Intensity Filter for Tracking Multiple Extended Objects in Clutter". In: *Proceedings of the 19th International Conference on Information Fusion (Fusion 2016)*. July 2016, pp. 1477–1484.

[97] Karl Granström, Stefphan Renter, Maryam Fatemi, and Lennart Svensson. "Pedestrian Tracking using Velodyne Data - Stochastic Optimization for Extended Object Tracking". In: *2017 IEEE Intelligent Vehicles Symposium (IV)*. 2017, pp. 39–46. DOI: 10.1109/IVS.2017.7995696.

[98] Karl Granström, Lennart Svensson, Stephan Reuter, Yuxuan Xia, and Maryam Fatemi. "Likelihood-Based Data Association for Extended Object Tracking Using Sampling Methods". In: *IEEE Transactions on Intelligent Vehicles* 3.1 (2018), pp. 30–45. ISSN: 2379-8858. DOI: 10.1109/TIV.2017.2788184.

[99] Biruk K. Habtemariam, Ratnasigham Tharmarasa, Thia Kirubarajan, Douglas Grimmett, and Cherry Wakayama. "Multiple Detection Probabilistic Data Association Filter for Multistatic Target Tracking". In: *14th International Conference on Information Fusion*. 2011, pp. 1–6.

[100] Robin Schubert, Christian Adam, Eric Richter, Sven Bauer, Holger Lietz, and Gerd Wanielik. "Generalized Probabilistic Data Association for Vehicle Tracking under Clutter". In: *2012 Intelligent Vehicles Symposium*. June, 2012.

[101] Michael Schuster, Johannes Reuter, and Gerd Wanielik. "Probabilistic Data Association for Tracking ExtendedTargets under Clutter using Random Matrices". In: *Proceedings of the 18th International Conference on Information Fusion (Fusion 2015)*. Washington, DC, USA, July 2015, pp. 961–968.

[102] Florian Meyer and Moe Z. Win. "Data Association for Tracking Extended Targets". In: *MILCOM 2019 - 2019 IEEE Military Communications Conference (MILCOM)*. 2019. DOI: 10.1109/MILCOM47813.2019.9020858.

[103] Xu Tang, Mingyan Li, Ratnasingham Tharmarasa, and Thiagalingam Kirubarajan. "Seamless Tracking of Apparent Point and Extended Targets Using Gaussian Process PMHT". In: *IEEE Transactions on Signal Processing* 67.18 (2019), pp. 4825–4838.

[104] Monika Wieneke and Wolfgang Koch. "A PMHT Approach for Extended Objects and Object Groups". In: *IEEE Transactions on Aerospace and Electronic Systems* 48.3 (July 2012), pp. 2349–2370. ISSN: 0018-9251. DOI: 10.1109/TAES.2012.6237596.

[105] Yaakov Bar-Shalom, Kuo-Chu Chang, and Henk A. P. Blom. "Tracking a Maneuvering Target using Input Estimation versus the Interacting Multiple Model Algorithm". In: *IEEE Transactions on Aerospace and Electronic Systems* 25.2 (1989), pp. 296–300.

[106] Yaakov Bar-Shalom, Peter K. Willett, and Xin Tian. *Tracking and Data Fusion: A Handbook of Algorithms*. YBS Publishing, 2011. ISBN: 9780964831278.

[107] Songhwai Oh, S. Russell, and S. Sastry. "Markov Chain Monte Carlo Data Association for Multi-Target Tracking". In: *IEEE Transactions on Automatic Control* 54.3 (2009), pp. 481–497. ISSN: 0018-9286. DOI: 10.1109/TAC.2009.2012975.

[108] Branko Ristic, Daniel Clark, Ba-Ngu. Vo, and Ba-Tuong Vo. "Adaptive Target Birth Intensity for PHD and CPHD Filters". In: *IEEE Transactions on Aerospace and Electronic Systems* 48.2 (2012), pp. 1656–1668. ISSN: 2371-9877. DOI: 10.1109/TAES.2012.6178085.