

**Extended and Coordinated Targets:
Tracker Design
Consideration and Approaches**

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Hosam Alqaderi
aus
Damaskus

Bonn 2025

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter/Betreuer: Prof. Dr. Wolfgang Koch

Gutachter: Prof. Dr. Peter Martini

Tag der Promotion: 27.10.2025

Erscheinungsjahr: 2025

This dissertation is dedicated to my parents, Salwa and Munthea, whose unwavering support, love, and encouragement have been the foundation of all my achievements. Your belief in me has been my greatest source of strength, and I owe everything I am to your endless sacrifices.

To my siblings, Alaa, Rama, and Mohammed, thank you for your constant encouragement and belief in my potential. Your love has been a guiding light throughout my life, and I am forever grateful.

To my fiancée, Sophie, thank you for your boundless patience, understanding, and love throughout this journey. Your unwavering support has been my anchor in the storm, and I could not have completed this work without you by my side.

I also extend my deepest gratitude to Sophie's family, whose warmth and kindness have made me feel like a part of your family. Your support has meant more to me than words can express.

This work is a testament to the love and support you have all given me. I am forever grateful.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Hosam Alqaderi

October 2025

Acknowledgements

Every time I visit Bonn and walk through its streets, I feel deeply grateful for the city and the country that have welcomed and embraced me. Germany, and Bonn in particular, have not only offered me an exceptional education and enriched my life but have also introduced me to many remarkable individuals who have significantly shaped my personal and professional journey. I am still eager to find a way to express my heartfelt appreciation to these people, whose influence has not only advanced my career but also broadened my perspective on life as a whole. Prof. Dr. Wolfgang Koch has been a major reason for my interest in this field. After his remarkable introduction to Sensor Data Fusion, I realized that I wanted to learn more about this subject and be as close to Prof. Koch as possible to gain all the knowledge I could. To this day, whenever I have the chance to attend one of Prof. Koch's talks, I find it both enjoyable and enlightening. His expertise and knowledge have consistently motivated me, and I believe this will continue to inspire me in the years to come.

Dr. Felix Govaers. From the moment I attended his lecture, I was profoundly inspired by the clarity and fervour with which he explained mathematical concepts. His ability to make complex formulas accessible and engaging made me think, "I want to talk and explain these ideas with the same skill and enthusiasm." Reflecting on my journey since then, I believe I have made significant strides toward achieving that aspiration, thanks in no small part to his influence and example.

Although not directly involved in my academic work, I want to express my deepest gratitude to someone whose influence has extended far beyond the confines of academia. Thank you, Prof. Dr. Susann Fiedler, for being a constant source of support when I needed it most. With her warm heart and brilliant mind, Dr. Fiedler has an extraordinary ability to inspire those around her, not only through her intellect but also through her kindness and empathy. Her unwavering support has provided invaluable guidance during times of uncertainty and offered encouragement when it was most needed. Koch, Govaers and Fiedler all have the ability to combine deep intellectual rigour with a genuine warmth of heart is rare, and I consider myself incredibly fortunate to have these people in during a pivotal time in my life.

I am fortunate to have had the privilege of being also under the supervision of Prof. Dr. Peter Martini. Lastly, I want to extend my sincere thanks to the Fraunhofer Institute and the University of Bonn for providing exceptional staff, cutting-edge research opportunities, and comprehensive support essential for academic and scientific success. I also wish to conclude this acknowledgements by expressing my gratitude to the city of Bonn and to Germany for offering me this incredible opportunity in life.

Abstract

Extended and coordinated target tracking involves advanced techniques for monitoring objects in dynamic environments, addressing challenges such as sensor resolution, the point target assumption and target behaviour.

Sensor resolution is crucial, as higher resolution enables more precise detection and tracking, particularly in complex scenarios where targets may be closely spaced or vary in size and shape. The point target assumption, commonly used in traditional tracking, simplifies targets as single points. However, in extended target tracking, targets are modeled with more complex shapes or patterns, accounting for variations over time.

Extended Target Tracking (ETT) becomes particularly important in scenarios where groups of targets move together, making it challenging to assign measurements to individual targets due to sensor resolution limitations and measurement noise. Even if the assignment problem is resolved, tracking each target individually within a group can be resource-intensive and impractical. To address this, Extended Target Tracking models the group as a single entity with shared kinematics, while simultaneously estimating the extent and shape of the group. This approach not only reduces computational demands but also enhances situational awareness. Examples of such scenarios include tracking fleets of aircraft or boats, formations of objects, swarms of drones, or groups of pedestrians.

Also, ETT plays a vital role in Advanced Driver-Assistance Systems (ADAS), addressing challenges when a single object spans multiple sensor resolution cells. These systems require precise information about the positions, sizes, and shapes of surrounding objects to ensure accurate perception. ETT improves safety and performance, enabling vehicles to avoid collisions, execute safe manoeuvres, and navigate complex environments, particularly when dealing with large or irregularly shaped objects in their surroundings.

Group targets, such as fleets of aircraft, boats, or swarms of drones, present unique challenges when operating autonomously and in a coordinated manner. Advances in distributed and multi-agent systems have enabled these entities to collaborate at scales surpassing human coordination, achieving complex objectives. However, these

innovations also complicate surveillance and defence efforts. Leveraging technologies like AI, communication, and computation, such systems achieve seamless coordination in dynamic environments, making them increasingly sophisticated and harder to counter.

Coordinated Target Tracking (CTT) addresses the limitations of traditional surveillance systems by improving situational awareness, optimizing resource management, and enhancing threat intelligence. CTT provides operators with accurate risk assessments, enabling informed and effective decision-making. This approach is critical in applications such as surveillance and defence, where identifying coordination and evaluating threats are essential.

In this dissertation, we will explore various approaches for simultaneously tracking both target shapes and kinematics by modeling target shapes as ellipses, particularly for LiDAR sensor measurements, and deriving a Gaussian mixture-based likelihood for improved accuracy. It introduces a novel Bayesian Gamma filter to estimate shape parameters, solving issues with negative estimates caused by Gaussian distributions. The approach extends to multivariate cases using the Wishart distribution, enabling robust and accurate shape parameter estimation. Unlike traditional Random Matrix (RM) methods, this work uses RM elements to represent shape extents, offering greater flexibility in modeling irregular shapes. These contributions enhance target tracking accuracy and reliability in dynamic environments.

Additionally, we will discuss methodologies to enhance tracking systems when multiple agents or targets exhibit coordinated behaviour, distinguishing coordinated targets from extended targets and examining multiple forms of coordination beyond traditional group target scenarios. The use of accumulated state densities (ASDs) is introduced to identify coordination over time, improving the separability and clustering of group targets. A novel method leveraging velocity vector geometry infers the degree of correlation, demonstrating that targets moving toward a common point are statistically closer. This approach relaxes the typical assumption that targets must be closely spaced, enhancing the detection of coordinated behaviours in dynamic environments. These innovations significantly advance the understanding and tracking of multi-agent systems.

Table of contents

List of figures	xv
List of tables	xix
Nomenclature	xxi
1 Introduction	1
1.1 Challenges	2
1.2 Main contributions	4
1.3 Thesis Outline	6
2 Sensors and Target Characteristics	7
2.1 Sensor Technology	7
2.1.1 Range-Angle Sensors	8
2.1.2 Local Polar to Local Cartesian Coordinates Conversion	9
2.2 Target Characteristics	12
3 Background on Bayesian Estimation	15
3.1 Overview of the Estimation Problem	15
3.1.1 Key Components and Modelling Approaches	15
3.2 Bayesian Estimation	19
3.2.1 Minimum Mean Square Error Estimator	20
3.2.2 Linear Minimum Mean Square Error Estimator	21
3.2.3 MMSE Estimator for Multivariate Gaussian	22
3.2.4 Unbiased Estimators	23
3.3 Dynamic Estimation for Target Tracking	24
3.3.1 Kalman Filter vs the Fundamental Equations of Linear Estimation	26

4	Extended Target Tracking Bayesian Framework	29
4.1	Overview of extended target tracking	29
4.1.1	Bayesian formalism	31
4.2	Extended Target Tracking: Spatial Elliptical Model for Laser Measure- ments	34
4.2.1	Introduction	34
4.2.2	Target Extent Model	36
4.2.3	Measurement Update	39
4.2.4	Prediction	41
4.2.5	Results	42
4.2.6	Conclusion	45
4.3	Extended Target Tracking: With Positive Support Density Filtering . .	46
4.3.1	Introduction	46
4.3.2	Uni-variate Gamma	50
4.3.3	Multivariate Gamma and Wishart Distribution	58
4.3.4	Conclusion	75
4.3.5	Summary	77
5	Coordinated Targets	81
5.1	Introduction	81
5.2	Background on Accumulated State Densities	83
5.2.1	Notion of ASD	84
5.2.2	ASDs Closed-form Representation	84
5.3	Accumulated State Densities Filter for Better Separability of Group- Targets	88
5.3.1	Spatial Clustering Method	89
5.4	Target State Separability	89
5.4.1	Simulation and Results	91
5.5	Automatic Identification of Coordinated Targets	94
5.5.1	Problem Statement	95
5.5.2	Probability of a group structure	96
5.5.3	Simulation and Results	100
5.6	Conclusion	102
6	Dissertation Conclusion	103
	References	105

Appendix A	109
A.1 List of Publications	109
Appendix B	111
B.1 Multiplication of Wishart density functions	111
B.2 Unbiased Wishart filter	111
B.3 On Self-Occlusion & Measurement Likelihood of Symmetric Shape . . .	113

List of figures

2.1	Demonstrates how various applications require different types of radar systems.	9
2.2	An example of Point Cloud created by Ibeo (MicroVision) Lidar's sensor. 10	
2.3	Spherical coordinate convention	12
4.1	Depicts a marine scenario featuring simulated radar detections from a marine surveillance radar. It also shows an extended target track with a shape moulded as an ellipse.	30
4.2	Illustrate possible measurement sources of an elliptical model for the shape of a LiDAR sensor located on the left side of the object in the left figure, and a radar located on the left side of the object in the right figure. 33	
4.3	Illustrates a set of noisy measurements resulting from measurement sources lying on the contour of an elliptical target.	35
4.4	Illustrates an elliptical target rotated by an angle α and a noisy measurement $\mathbf{z}$ originating from a particular measurement source $\mathbf{y}$ lying on the contour of the ellipse at an angle θ	37
4.5	This figure shows a heat map representing the probability density of the Gaussian mixture with the number of components corresponding to the measurement sources shown by the black points	39
4.6	The figure illustrates the evaluation results of the proposed approach as shown by the blue curve compared to the Random Matrix approach shown by the green curve.	44
4.7	In this simulation result, the single extended target ground truth is represented by the black rectangle and the dark blue points show the measurements reflected from the target to a LiDAR sensor located at coordinate $(0,0)$. The estimated target shape by proposed approach is represented by the red ellipse, while the Random Matrix-Feldmann approach is represented by the blue ellipse.	45

4.8	Exemplary distributions of a parameter $d = 2.1$ with a standard deviation $\sigma = 4$. The Gaussian density (black) has a significant part of its probability mass within the negative real axis whereas the Gamma distribution is limited to the positive real axis.	47
4.9	Example plots of the Gamma distributions for different parameters, where $\theta = 1$ is fixed.	51
4.10	Plots of the Gamma distributions for the exact and unbiased update rules with the parameters from Table 4.1	55
4.11	Different snapshots are taken from a simulation of Ibeo automotive LiDAR in an urban scenario: The white frame box represents the extent of the obtained tracks using the proposed Gamma filter approach. The orange points show the LiDAR point cloud relative to the red vehicle (the ego vehicle).	56
4.12	The <i>Intersection over Union</i> results of the shape tracking using the proposed Gamma filter as shown by the blue curve which corresponds to track with id zero in figure 4.11 (a) and (b). Also by the red curve which corresponds to track id one in figure 4.11 (c) and (d).	57
4.13	This figure illustrates the <i>Root Mean Square Error</i> results of the estimated target's length and width. Where the track's number in the figure's legend corresponds to the track's id in figure 4.11.	57
4.14	Photogrammetric technique is used to take a spatial measurement from photographs to measure the pixels' positions related to the selected features, in this case, the edges of the object.	60
4.15	Two measurement sources $\mathbf{H}^i r^i$ and $-\mathbf{H}^i r^i$ correspond to a single segment at angle θ^i of a symmetric star-convex shape. Measurements stemming from any of the two sources are associated with the segment in red. . .	65
4.16	Illustration of the estimated shapes by Wishart filter, using general and independent measurement models.	72
4.17	In this simulation result, the extended target ground truth extent is represented by the black dashed shape and the dark blue point measurements resemble the LiDAR sensor reflection of one side of the target contours. The red and the yellow solid curves represent the estimated target extent by the proposed approach, corresponding to the independent and SE models. The GP approach estimated extent is represented by the blue solid curve.	73

4.18	The figure illustrates the IoU results of the proposed Wishart filter for both independent and SE models shown in the red and yellow curves respectively. It is also compared to the GP approach shown by blue curve.	76
5.1	Illustration of different coordination structures.	82
5.2	Schematic comparison of ASD and Kalman filter fusion. Image courtesy of Govaers and Koch [17]	83
5.3	This figure shows the Mahalanobis distance obtained by 1000 Monte Carlo run of targets tracked using different state density sizes, where the total number of tracking frames is 200. The blue colour corresponds to targets separated by $50m$, while the orange colour corresponds to targets separated by $65m$. The error bar correspond to one standard deviation of the Mahalanobis distance.	90
5.4	In this simulation scenario, two group-targets and three point-targets are used. The blue dashed line shows the retrodicted states of the targets correspond to time frame $k - 1$	92
5.5	The figure illustrates the identified group-targets and point-targets by DBSCAN using $ASDF_{n=20}$, which are coded by color at a different simulation frames.	92
5.6	The figure shows the average number of wrong clusters of DBSCAN with Kalman filter comparing to DBSCAN with ASDF using different time window sizes.	93
5.7	Illustration of the group structure $G_t = [1, 1, 2, 2, 3]^T$ formed by 5 targets.	95
5.8	The heading of targets i and j shows that both targets are moving toward a common point. The radial velocity components resulting from projecting the velocity vector on the vector connecting i and j are shown in blue.	99
5.9	Target i and j are spread by a fixed distance. The heading of target i is also fixed to about 45 degrees, while target j 's heading rotates along the circle, Also, target i 's velocity is greater than target j 's velocity. The orange curve shows the different log-likelihood of targets i and j to belong to the same group for each different heading of j . The log-likelihood is obtained by taking the log of $f(\mathbf{x}_{k:n}^i, \mathbf{x}_{k:n}^j)$ defined in (5.34).	100
5.10	Illustrates the most likely group structure at different frames which are represented by colour-coded groups and independent targets.	101

-
- B.1 This figure compares the approximation of the measurement likelihood by a uni-variate Gaussian with zero-mean in two scenarios. Scenario (a), represents the case where two ends of the segments are observable, while (b) represents the case when only one end is observable. The measurement likelihood is shown by the blue solid curve. While the red solid curve represents the approximation of the blue curves by a zero-mean Gaussian. 114
- B.2 In this figure the measurements samples of two segments with radius $l = 1$ are illustrated by the blue points in two scenarios: (a) both ends of the segments are observable and (b) a single end is observable. Also, the figure compares the measurement likelihood obtained by the general model (4.98) and the independent model (4.99) illustrated by the black and red dashed ellipses respectively. 115

List of tables

4.1	Example for the Bayes update for a given prior and likelihood using the exact and the unbiased rules.	54
4.2	Root-mean-square error of different point estimators are listed in the table. The numbers are averaged over 1000 MC runs. The parameters n and $\mathbf{C}$ are calculated as in (B.6) and (B.7).	70

Nomenclature

List of Symbols

$\mathbf{a}, \mathbf{b}, \dots, \mathbf{x}, \mathbf{y}, \mathbf{z}$ Vectors

$\mathbb{C}[\dots]$ Covariance matrix

$|\dots|$ Determinant of a matrix, norm of a vector

$\mathbb{E}[\dots]$ Expectation

$(\dots)^{-1}$ Inverse

n_z Number of measurements

φ Azimuth/bearing angle

$\mathbb{R}$ Set of real numbers

t Time

$tr(\dots)$ Trace

$(\dots)^\top$ Transpose

$\mathbf{u}_k$ White Gaussian noise at time t_k

$\mathbf{A}, \mathbf{B}, \dots, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$ Matrices

$\mathbf{x}_k$ State vector at time t_k

$\mathcal{Y}_k$ Set of measurement sources at time t_k

$\mathbf{z}_k$ Measurement vector at time t_k

$\mathcal{Z}_k$ A set of measurements at time t_k

$\mathcal{Z}^k$ Time series of measurements up to time t_k

Acronyms / Abbreviations

ADAS Advanced Driver-Assistance Systems

ASD Accumulated State Density

ASDF Accumulated State Density filte

BLUE Best Linear Unbiased Estimator

CTT Coordinated Target Tracking

DBSCAN Density-based Spatial Clustering of Applications with Noise

EKF Extended Kalman Filter

ETT Extended Target Tracking

GGIW-PHD Gamma Gaussian Inverse Wishart PHD

iid Independent and identically distributed

IoT Internet of Things

LiDAR Light Detection and Ranging

LMMSE linear minimum mean square error estimator

LMSE Least Mean Squares Estimator

LSE Least Squares Estimator

METT Multiple Extended Target Tracking

MLE Maximum Likelihood Estimator

MMSE Minimum Mean Square Error Estimator

MSE Mean Squared Error

PDF Probability Density Function

PHD Probability Hypothesis Density

PMF Probability Mass Function

Radar Radio Detection And Ranging

RHM Random Hypersurface Model

RM Random Matrix

WLS Weighted Lease Squares Estimator

Chapter 1

Introduction

The world is changing rapidly—a statement that has held true throughout history. In recent decades, however, the most significant changes have been driven by remarkable advancements in processing power and computational technology. We now have more computational capability than ever before, accelerating progress in various fields, from smart home and wearable devices to space exploration. These developments are transforming the way we live, work, and explore the universe.

Driven by our optimistic dreams and imagination, many of us believe that one day, life will be like this: After waking up in your smart home, you step out and head to your vehicle, powered by green energy. As you approach, you simply tell it to open the door and start the engine. The vehicle recognises you instantly, thanks to advanced voice and facial recognition systems. Once inside, you'll notice there's no driver's seat because the vehicle drives itself. You just tell it where to go. This vehicle might even have a built-in coffee machine or a small robot that prepares your breakfast. Thanks to the Internet of Things (IoT), it automatically orders more coffee beans and your favourite breakfast items when they're running low. It knows exactly what you like because you'll be surrounded by sensors connected to a system that knows you better than you know yourself. If you happen to be living in Germany at that time, you'll probably be handed a nice, thick document about data protection ("Datenschutz") to sign. And of course, you'll eagerly sign it without a second thought because, after all, you've likely been bombarded with ads promising something like: "We'll create a system that anticipates your needs, crafting a seamless and personalised experience just for you." Who could resist that?

Well, you've arrived at your destination, and as your vehicle automatically begins recharging its green energy source, you step out and look up. The sky is filled with drones, each on its own mission. Thanks to advanced multi-agent systems and

lightweight, high-resolution sensors, these drones can coordinate effortlessly. They deliver your packages, clean the solar panels on your roof, and even help find your lost dog after you took him for a walk in some artificial forest.

Indeed, every part of this story revolves around sensors. It's hard to imagine life and future without systems equipped with rapidly advancing sensors. This progress also requires corresponding improvements in algorithms. We need algorithms that are not only accurate and based on mathematical or explainable data-driven models but also fast and precise, outperforming human performance. For example, sensors in vehicles must perceive their surroundings, understand the objects around them, and predict their movements. These sensors are expected to see further and better than humans. Similarly, a comprehensive surveillance system is essential for our airspace, especially given the new challenges posed by the increasing number of small autonomous flying objects, such as drones. Drones might enter restricted areas, such as crowded spaces or near airports, potentially conflicting with flight paths and posing a threat to public safety. Additionally, drones can be used for illegal activities such as drug smuggling, espionage, and weaponization. Therefore, we need systems capable of tracking drones, detecting suspicious behaviours, and identifying the coordination among them.

Surveillance systems that track targets in dynamic environments are crucial across various fields, including defence, surveillance, environmental monitoring, autonomous systems and robotics. As these applications advance, the demands on tracking systems increase, requiring them to handle increasingly complex scenarios. This dissertation focuses on two key concepts in this area: extended target tracking and coordinated target tracking. Extended target tracking deals with targets that occupy more than a single point in space, necessitating models that account for their shape, size, and dynamic behaviour. In contrast, coordinated target tracking involves monitoring multiple targets that exhibit interrelated or synchronised behaviours. Both approaches present significant challenges but are essential for enhancing situational awareness and decision-making in complex environments.

1.1 Challenges

Traditional tracking systems often rely on the point target assumption, which simplifies targets as single points. While this assumption works for small or distant targets, it becomes inadequate when dealing with extended targets whose physical dimensions and shapes must be considered. Additionally, many tracking systems are designed to track individual targets in isolation, ignoring the complexities introduced by coordinated

behaviour among multiple targets. These limitations can lead to inaccuracies, especially in environments where targets may overlap, manoeuvre unpredictably, or exhibit coordinated movement. There is a clear need for advanced tracking methodologies that can accurately and reliably track both extended and coordinated targets in complex scenarios. The main challenges in *single extended target tracking* can be summarised as follows:

- **Sensor Resolution:** High-resolution sensors are essential for capturing detailed information about target shapes and behaviours. However, they also introduce challenges related to data processing, management, and data association. This challenge is typically addressed by clustering the reported measurements from the sensors at each frame as a preprocessing step. Alternatively, the problem can be tackled at the association stage of the tracking pipeline, where the typical assumption that each target generates at most one detection is relaxed.
- **Modeling Complexity:** Extended target tracking requires accurate modeling of the target's extent (shape). These models can range in complexity from simple geometric shapes, such as lines, circles, or ellipses, to more sophisticated random, symmetric, and asymmetric forms. Depending on the target shape model, it is necessary to derive a measurement model that maps the target state—encompassing both kinematics and shape—to the measurement state space. When the target state is augmented with shape parameters, these models often become highly nonlinear, adding to the complexity of the tracking process.

Different challenges arise when multiple targets exhibit some form of coordination, and these can be summarized as follows:

- **Definition of Coordination:** The term "coordinated targets" is somewhat ambiguous. One might argue that closely spaced targets flying in a group represent a form of coordination. However, members of a coordinated group can also spread out to attack a target from different angles, ensuring they don't overlap their efforts, or they might separate widely to search a broader area. Additionally, coordinated targets can have different topologies, such as having a group leader or operating independently to achieve a common goal. There might be explicit communication between the targets, or the coordination might be implicit, where each target observes the actions of other members to form a strategy for the next move. This challenge is typically addressed by first defining the specific form of coordination that the surveillance system aims to detect.

- **Number of Possible Groups:** In a surveillance system with multiple targets, determining which targets are coordinating and identifying the most likely group structure can be challenging. The space of possible group structures is vast because the number of possible configurations corresponds to the number of ways to partition a set of targets, which is given by the Bell number. This means that as the number of targets increases, the number of potential groupings grows exponentially, making the problem increasingly complex.

1.2 Main contributions

This dissertation contributes to the fields of Extended Target Tracking (ETT) and Coordinated Target Tracking (CTT) by addressing several key challenges.

Contributions to the Field of Extended Target Tracking

One of the main challenges in ETT is the complexity of modeling target shapes and their corresponding measurement likelihood functions. In this dissertation, we demonstrate how a target's shape can be modeled as an ellipse, using a measurement likelihood model where the measurement density is concentrated around the contour of the target shape, particularly for measurements obtained from Light Detection and Ranging (LiDAR) sensors. Although elliptical shapes are commonly used in target tracking literature, most elliptical models assume radar-type measurements, where the measurement density is represented by a Gaussian distribution centred at the middle of the target's extent. Our contribution in this area is the derivation of an explicit mathematical function for the measurement model, derived directly from the standard geometrical function of the elliptical shape. The measurement likelihood approximated in a form of a Gaussian mixture, is computed and incorporated into the Extended Kalman Filter (EKF) framework. This approach provides a more accurate representation of measurements for sensors like LiDAR, enhancing the effectiveness of extended target tracking.

Another contribution of this dissertation in the same area is the development of a robust novel approach for estimating shape parameters that does not rely on the typical assumption of modeling these parameters as Gaussian random variables. The motivation for this work was to address the issue of negative shape estimates, which has been reported in the literature. The root of this problem lies in the fact that the Gaussian distribution has negative support, meaning it can produce negative values.

When the estimated parameter is relatively small compared to the variance, a significant portion of the Gaussian density falls into the negative range.

To solve this issue, we propose using a distribution with positive support only—namely, the Gamma distribution. We have developed a Bayesian Gamma filter to estimate quantities that are known to take only positive values. We derive the exact update formula for the Gamma distribution and demonstrate how the prediction step can be performed using this approach. This method effectively eliminates the problem of negative shape estimates, providing a more accurate and reliable estimation filter. Moreover, we generalize this approach to the multivariate Gamma distribution, particularly to a special case known as the Wishart distribution. Similarly, we derive a Bayesian Wishart filter, demonstrating the exact update equation and prediction process. In this approach, the shape is represented by a Random Matrix (RM) with dimensions that can be chosen based on the level of detail required for the shape.

When we mention that the shape is represented by an RM, those familiar with the ETT literature might immediately think of the pioneering Random Matrix approach by Koch [26]. However, our approach differs slightly. While Koch’s method estimates the spread of measurements as an estimation of a covariance matrix of Normally distributed measurements, in our approach, the elements of the RM represent random variables that contain information about the shape’s extent in specific directions. This allows for the estimation of random, irregular shapes with greater flexibility.

Contributions to the Field of Coordinated Targets Tracking

Coordinated targets represent a relatively unexplored area of research, with most existing studies focusing primarily on one form of coordination, where coordinated targets move in a group (group targets). This form of coordination is often treated as an extended target tracking problem. In this dissertation, we distinguish between the characteristics of extended and coordinated targets. While group targets are one form of coordination, we explore other forms and define a common characteristic: members of a coordinated group are assumed to converge toward a common point, relaxing the assumption that targets are closely spaced, as is typical in traditional group target scenarios.

Detecting coordinated behaviour or a multi-agent system is nearly impossible with a single frame; multiple frames (past information) are required to identify coordination. We show how accumulated state densities (ASDs) provide an elegant framework that can be utilised to enhance the identification of different forms of coordination. For

the group target form of coordination, we demonstrate how ASDs can improve the separability of groups, leading to more accurate clustering of group targets.

Beyond group coordination, we proposed an approach that leverages the geometry of targets' velocity vectors to infer the degree of correlation over a chosen ASD sliding time window. This method reveals that coordinated or correlated targets appear statistically closer to one another, meaning that target pairs moving toward a common point have a higher likelihood of belonging to the same group compared to other targets. This approach opens new avenues for computing degree of correlation based on features of targets state vector and identifying coordinated behaviours in dynamic environments.

1.3 Thesis Outline

This dissertation is divided into three main parts. The first part, covered in Chapters 2 and 3, provides essential background information. Chapter 2 discusses the various types of sensors and targets, along with their key characteristics. Chapter 3 reviews foundational concepts of the estimation problem, with a particular focus on Bayesian estimation principles. These chapters lay the groundwork for the more advanced topics explored in the subsequent sections of the dissertation.

The second part of this dissertation, covered in Chapter 4, discusses the problem of extended target tracking and presents various approaches to address it. Section 4.2 primarily focuses on the Elliptical Model for Extended Target Tracking Using Laser Measurements. In Section 4.3, an alternative approach is proposed, highlighting the limitations of modeling target size as Gaussian distributions. Here, we introduce the Gamma distribution as an alternative, which has positive support only. This section also generalises the concept to a multivariate form, demonstrating how it can be achieved using a special case of the multivariate Gamma distribution.

The third part of this dissertation is covered in Chapter 5. It begins with an introduction to the problem of coordinated targets, followed by background on the Accumulated State Densities concept, which forms the basis of the approaches introduced in this chapter. In Section 5.4, target state separability is discussed, and an improvement that enhances the clustering of group target types is introduced. Finally, a more sophisticated approach aimed at determining the most likely group structure that can be formed by a set of targets is proposed in Section 5.5.

Chapter 2

Sensors and Target Characteristics

For the successful design of a sensor data fusion system, it is essential to have reliable and sufficient knowledge of sensor performance. This includes an understanding of the sensor's accuracy, precision, sensitivity, and resolution, as well as its limitations and potential sources of error.

In addition to understanding individual sensor characteristics, it is important to consider the characteristics of the target for which the system is designed to track, as different applications require different types of sensor models and varying levels of sensor performance. By carefully selecting and integrating the right sensors and developing appropriate data fusion algorithms, it is possible to create a highly effective sensor system that provides valuable insights and improves decision-making across a wide range of fields.

2.1 Sensor Technology

Our environment is full of physical properties that change over time. Sensors are instruments designed to measure and quantify various parameters in the environment and convert them into readable outputs, such as electrical signals. They can measure different parameters—physical, chemical, or biological—based on their operating principles. Progress in sensor technology has led to the development of numerous novel applications, including wearable devices, smart cities, home automation, and self-driving cars. Additionally, sensor technology plays a crucial role in defence applications by enabling situational awareness. For example, various types of sensors, such as radar, sonar, infrared, LiDAR, and electro-optical sensors, are used to detect and track targets and identify potential threats on land, at sea, or in the air.

2.1.1 Range-Angle Sensors

Sensors that can measure both the range and angle to an object of interest are often referred to as "range-angle sensors" and may also be called spatial sensors. Target tracking is a very common application for range-angle sensors, and several technologies fall under this category, including Radar (Radio Detection and Ranging), LiDAR (Light Detection and Ranging), Sonar (Sound Navigation and Ranging), and Ultrasonic sensors. Since the work presented in this dissertation pertains specifically to Radar and LiDAR sensor types, our discussion will be restricted to these two types.

RADAR

The basic principle of radar operation is straightforward: transmitting pulses of energy in the form of electromagnetic radiation, particularly radio waves at microwave frequencies. In the presence of an object, these transmitted pulses are reflected back and then picked up by receiving antennas. The received energy, reflected from the object, is processed to determine the presence of the object and measure various kinematic quantities. Since electromagnetic energy travels through air at a constant speed, the radar system can determine the distance to the reflecting object by measuring the time it takes for the transmitted pulses to travel. In addition to measuring object range, modern radar systems are capable of measuring additional kinematic quantities, such as object angle (azimuth and elevation) and range rate. Radar has numerous applications in different fields, including aviation, defence, oceanography, geological surveys, and automotive industries. Fig. 2.1 illustrates different types of radar systems used for various applications.

LiDAR

As laser range technology continues to enable a broad range of applications, such as localisation, target tracking, and 2D and 3D mapping, LiDAR sensors are becoming increasingly important. In principle, the basic operation of LiDAR is similar to that of radar, as both sensors emit energy towards an object and measure the time it takes for the energy to bounce back to the sensor. However, LiDAR emits a laser beam instead of radio waves.

While the principle may be similar, LiDAR generates a distinct type of data known as Point Cloud, which is extremely dense, precise, and high-resolution data in both 2D and 3D. This type of data has enabled the advancement for a wide range of applications, including:



(a) air surveillance radar



(b) maritime radar



(c) airborne radar



(d) weather radar

Fig. 2.1 Demonstrates how various applications require different types of radar systems.

- Autonomous vehicles utilise point cloud to quickly and accurately create a detailed 3D map of the environment, allowing the car to make informed decisions about how to navigate around obstacles and other vehicles. An example of LiDAR data is shown in Fig 2.2.
- Create detailed maps of archaeological sites or geological formations. By analysing the point cloud data, researchers can gain insights into the history and composition of these sites.
- Create detailed topographical maps of the environment, which can be used for a variety of applications such as urban planning, civil engineering, and environmental studies.

Various laser sensors are available that generate distinct kinds of data, but this dissertation will focus only on a laser sensor known as Automotive LiDAR.

2.1.2 Local Polar to Local Cartesian Coordinates Conversion

Performing Euclidean mathematics and applying the laws of motion in Cartesian coordinates is more convenient. However, since Range-Angle Sensors typically re-

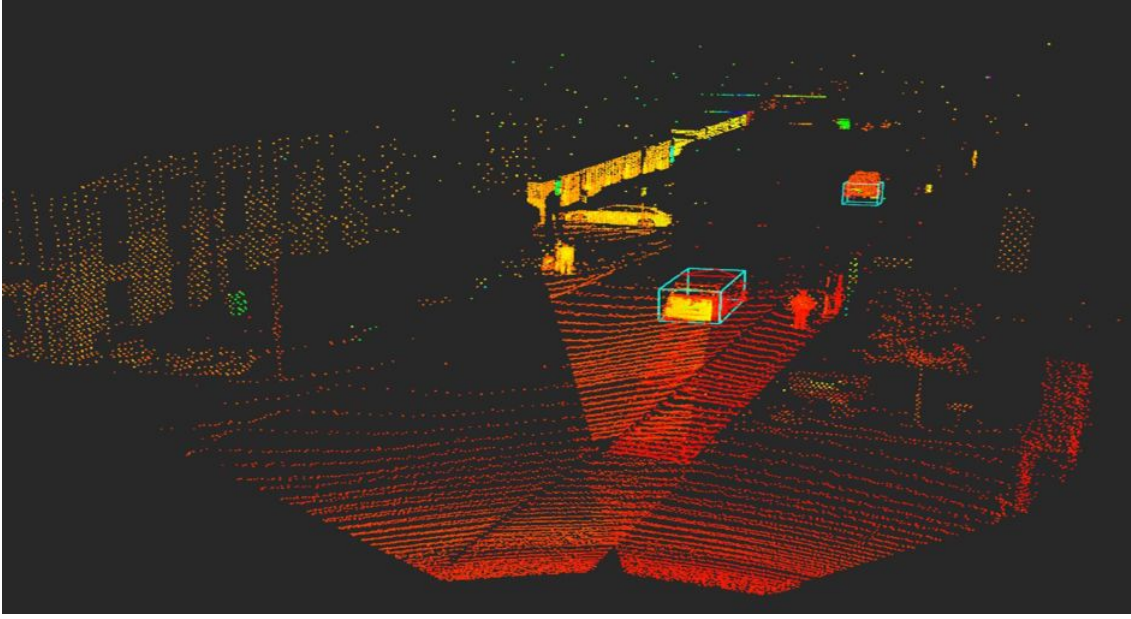


Fig. 2.2 An example of Point Cloud created by Ibeo (MicroVision) Lidar's sensor.

port measurements in polar coordinates, it is frequently necessary to convert these measurements into Cartesian coordinates.

2D Conversion

Given a range measurement r and azimuth angle, or bearing φ , the measurement vector is given in polar coordinates as

$$\mathbf{z}_{polar} = [r, \varphi]^T. \quad (2.1)$$

One can express the errors in range and bearing as a function of their respective values. The covariance of the polar measurement vector quantifies the uncertainty of the polar measurements and is commonly described by a diagonal matrix as:

$$\mathbb{C}[r, \varphi] = \begin{bmatrix} \sigma_r^2(r, \varphi) & 0 \\ 0 & \sigma_\varphi^2(r, \varphi) \end{bmatrix}. \quad (2.2)$$

The transformation described below can be used to derive the Cartesian measurement vector from the polar measurement vector

$$\mathbf{z}_{cartesian} = \begin{bmatrix} x \\ y \end{bmatrix} = f(r, \varphi) = \begin{bmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{bmatrix} \quad (2.3)$$

The covariance $\mathbb{C}[r, \varphi]$ can also be converted into Cartesian coordinates using first order approximation

$$\mathbb{C}[x, y] = J \mathbb{C}[r, \varphi] J^\top, \quad (2.4)$$

where J is the first-order partial derivatives of all vector-valued function element 2.3, and is often called the Jacobian matrix

$$J = \frac{\partial f(r, \varphi)}{\partial(r, \varphi)} = \begin{bmatrix} \cos(\varphi) & -r \sin(\varphi) \\ \sin(\varphi) & r \cos(\varphi) \end{bmatrix}. \quad (2.5)$$

3D Conversion

The measurement of range-angle sensor in 3D is expressed using spherical coordinates as

$$\mathbf{z} = [r, \varphi, \psi]^\top. \quad (2.6)$$

where r is the range, φ is the azimuth (or bearing) and ψ is the elevation. It should be noted that for compatibility with the notation, the azimuth angle is measured starting from the positive y-axis in clockwise direction and the elevation angle is measured from the azimuthal plane (the plane containing the x and the y axes) rather than the z axis, Fig 2.3.

One can express the errors in range, bearing and elevation as a function of their respective values. The spherical measurement vector's errors are commonly described by a diagonal matrix as

$$\mathbb{C}[r, \varphi, \psi] = \begin{bmatrix} \sigma_r^2(r, \varphi, \psi) & 0 & 0 \\ 0 & \sigma_\varphi^2(r, \varphi, \psi) & 0 \\ 0 & 0 & \sigma_\psi^2(r, \varphi, \psi) \end{bmatrix}. \quad (2.7)$$

The transformation described below can be used to derive the Cartesian measurement vector from the polar measurement vector

$$\mathbf{z} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = f(r, \varphi, \psi) = \begin{bmatrix} r \sin(\varphi) \cos(\psi) \\ r \cos(\varphi) \cos(\psi) \\ r \sin(\psi) \end{bmatrix} \quad (2.8)$$

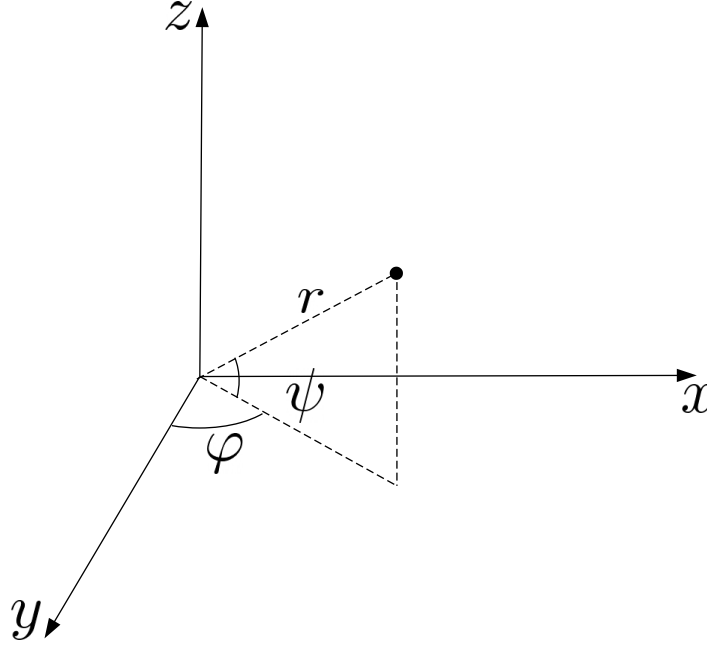


Fig. 2.3 Spherical coordinate convention

The covariance $\mathbb{C}[r, \varphi, \psi]$ can also be converted into Cartesian coordinates using first order approximation

$$\mathbb{C}[x, y, z] = J \mathbb{C}[r, \varphi, \psi] J^T, \quad (2.9)$$

where J is the first-order partial derivatives of all vector-valued function element 2.8, and is often called the Jacobian matrix

$$J = \frac{\partial f(r, \varphi, \psi)}{\partial (r, \varphi, \psi)} = \begin{bmatrix} \sin(\varphi) \cos(\psi) & r \cos(\varphi) \cos(\psi) & -r \sin(\varphi) \sin(\psi) \\ \cos(\varphi) \cos(\psi) & -r \sin(\varphi) \cos(\psi) & -r \cos(\varphi) \sin(\psi) \\ \sin(\psi) & 0 & r \cos(\psi) \end{bmatrix}. \quad (2.10)$$

2.2 Target Characteristics

In tracking problems, the physical, real-world objects of interest always have special characteristics. These characteristics might be, size, shape, manoeuvring capability, radar cross-section (RCS), infrared signature and other. To design a tracking system and data fusion algorithm, it is essential to first define and understand the characteristics

of the target of interest and the level of situational awareness required for certain safety systems or defence applications. The sensor properties need to be also taken into account, for example, the differences between extended target tracking and point target tracking are primarily influenced by the sensor properties, especially the sensor resolution, rather than object properties such as spatial extent. The relevant target characteristics identified in this dissertation are

- Point target:
At most a single measurement could stem from each target per time step, i.e., a single resolution cell is occupied by the target.
- Extended target:
Each target generates several measurements per time step. The target's measurement sources are spatially distributed across its rigid body, occupying multiple resolution cells. To address this, the various measurement sources are regarded as a single entity with common kinematics. To adequately construct the measurement model, it is necessary to estimate of the target's kinematics, as well as target's extent (its shape and size). An example of extended target are tracking ships or truck.
- Group target:
A group target consists of two or more closely spaced sub-targets that share a common kinematics. Thus, the group target occupies several resolution cells; each sub-targets may occupy either one or several resolution cells. Further, the targets are not tracked individually but are instead treated as a single entity. Similar to extended target tracking, estimation of the target's kinematics, as well as target's extent is required for group target tracking.
- Coordinated target:
The common assumption made in the both extended and group-target tracking problem is that the measurement sources are closely spaced and have common kinematics. However, this assumption doesn't hold for coordinated targets. A coordinated targets such as swam requires a set of rules which each member of the group follows to accomplish a shared goal. Members of the coordinated group can spread out and can be widely separated over the field of view. This characteristic is not directly related to the target physical characteristics neither to the sensor properties, but it is related to the targets' behaviour. Detecting coordinated targets is becoming increasingly important in achieving advanced

situational awareness and threat intelligence. Knowledge about group behaviour can significantly assist in target tracking; for example, if the group consistently moves in a certain direction or pattern, this data can be leveraged to predict the next location of the targets. Additionally, recognising collective behaviours can significantly enhance the prioritisation of effectors in defence systems.

Chapter 3

Background on Bayesian Estimation

3.1 Overview of the Estimation Problem

This chapter reviews the theory and fundamentals of Bayesian estimation techniques, which provide the foundation for the various approaches discussed in this dissertation. Before diving into *Bayesian estimation*, we will briefly discuss the estimation problem from a non-Bayesian perspective.

3.1.1 Key Components and Modelling Approaches

In general, the estimation problem can be summarised as the challenge of determining unknown quantities of interest based on a set of observed data. The goal is to make inferences or predictions about these unknowns using the available observations.

The observations can be either direct measurements of the unknown quantities or indirect ones. The essential elements of the estimation problem consist of three main components:

- **Unknowns:** the quantities or the variables we want to estimate. These quantities are usually a state $\mathbf{x} = [x^1, x^2, \dots, x^{n_x}]^\top \in \mathbb{R}^{n_x}$, or a parameters $\theta = [\theta^1, \theta^2, \dots, \theta^{n_\theta}]^\top \in \mathbb{R}^{n_\theta}$ of the system¹.
- **Measurement:** refers to the information we can observe about the unknowns. It is described by the measurement vector $\mathbf{z} = [z^1, z^2, \dots, z^{n_z}]^\top \in \mathbb{R}^{n_z}$

¹In some literature, parameter estimation is considered a **static estimation problem**, while state estimation is regarded as a **dynamic estimation problem**. In this work, we use the term "state estimation," which does not necessarily imply that the state is dynamic.

- **Measurement model:** a function which describes the relation between the measurement and the unknowns.

Deterministic model

The measurement model is a mathematical representation that describes the relationship between the physical world and the corresponding measurements. One common way to formulate the measurement equation is by assuming that the measurement vector is deterministic:

$$\mathbf{z} = \mathbf{H}\mathbf{x}, \quad (3.1)$$

where $\mathbf{H}$ is the design matrix that maps the state vector $\mathbf{x}$ to the measurement vector $\mathbf{z}$. In the case of a non-linear relationship, the equation can be expressed as:

$$\mathbf{z} = f(\mathbf{x}), \quad (3.2)$$

Here, $f(\cdot)$ is a non-linear function that maps the state vector $\mathbf{x}$ to the measurement vector $\mathbf{z}$.

Under the assumption that both the state vector and the measurement vector are deterministic, the unknown variables can be obtained by solving the linear system of equations. In the non-linear case, finding the inverse function $f^{-1}(\mathbf{z})$ may provide a solution if it exists. However, in real-world scenarios, measurements are often corrupted by noise, making it impossible to solve the linear system directly. To obtain an estimate in such cases, a common approach is to incorporate the measurement noise into the model by adding an error term, $\mathbf{u}$, to the measurement equation:

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{u}. \quad (3.3)$$

After incorporating the measurement noise into the model by adding an error term $\mathbf{u}$ to the measurement equation, one can employ an optimisation algorithm to find a solution for the system of equations by minimising a cost function. This optimisation process yields an estimator, which is a mathematical function of the measurement vector, and produces an estimate $\hat{\mathbf{x}}$ of the unknown variables $\mathbf{x}$,

$$\hat{\mathbf{x}} = f(\mathbf{z}) \quad (3.4)$$

A common algorithm that provides an estimate under the "minimum distance" cost function is the *Least Squares Estimator* (LSE). The objective of the LSE method is to

determine the parameter values $\mathbf{x}$ that minimise the sum of the squared differences between the observed data and the model's predicted values, as represented by the error term $\mathbf{u}$ in the equation (3.3).

$$\mathbf{x}_{LSE} = \arg \min_{\mathbf{x}} (\mathbf{u}^\top \mathbf{u}) \quad (3.5)$$

$$= \arg \min_{\mathbf{x}} (\mathbf{z} - \mathbf{H}\mathbf{x})^\top (\mathbf{z} - \mathbf{H}\mathbf{x}). \quad (3.6)$$

Finding a solution for this minimization problem involves taking the first derivative of the function and solving for the value(s) at which the derivative equals zero,

$$\nabla_{\mathbf{x}} (\mathbf{z} - \mathbf{H}\mathbf{x})^\top (\mathbf{z} - \mathbf{H}\mathbf{x}) = 0 \quad (3.7)$$

$$-2\mathbf{H}^\top \mathbf{z} + 2\mathbf{H}^\top \mathbf{H}\mathbf{x} = 0 \quad (3.8)$$

$$\mathbf{H}^\top \mathbf{H}\mathbf{x} = \mathbf{H}^\top \mathbf{z} \quad (3.9)$$

$$\mathbf{x}_{LSE} = (\mathbf{H}^\top \mathbf{H})^{-1} \mathbf{H}^\top \mathbf{z} \quad (3.10)$$

In many real-world scenarios, not all measurements are of the same quality. To account for these differences, one can assign higher weights to reliable measurements and lower weights to poorer ones. This can be achieved by incorporating a weight matrix $\mathbf{W}$ into equation 3.10, resulting in a type of estimator called the *Weighted Least Squares Estimator* (WLS),

$$\mathbf{x}_{WLS} = (\mathbf{H}^\top \mathbf{W}\mathbf{H})^{-1} \mathbf{H}^\top \mathbf{W}\mathbf{z} \quad (3.11)$$

Stochastic model

The LSE and WLS estimators do not account for the stochastic properties of the measurements. However, when dealing with a stochastic¹ model, it becomes necessary to derive an estimator that can provide optimal estimates, both in terms of precision and unbiasedness. To achieve this, it is essential to have knowledge of the statistical expectation and dispersion of the measurement vector, denoted as $\mathbb{E}[\mathbf{z}]$ and $\mathbb{C}[\mathbf{z}]$ respectively.

¹Stochastic models are often referred to as random models; thus, the terms "stochastic" and "random" are used interchangeably in this work.

Considering that, the measurement model $\mathbf{H}\mathbf{x}$ is the first moment of the measurement $\mathbf{z}$, and $\mathbf{Q}_u$ is the second central moment of $\mathbf{z}$.

$$\begin{aligned}\mathbb{E}[\mathbf{z}] &= \mathbf{H}\mathbf{x} \\ \mathbb{C}[\mathbf{z}] &= \mathbb{C}[\mathbf{u}] = \mathbf{Q}_u,\end{aligned}$$

Similar to the LSE, one can derive an estimator for the unknown random variable $\mathbf{x}$ that minimises a cost function. However, since the model is stochastic, it is necessary to adjust for both the varying variance of the errors across different values of the random variable and any correlated errors. This can be accomplished by weighting the measurements by the inverse of their corresponding covariance, $\mathbb{C}[\mathbf{z}]$. This method is known as the Best Linear Unbiased Estimator (BLUE), which has the same form as the Weighted Least Squares (WLS) estimator 3.11, except that the BLUE weights are based on the inverse of the estimated covariance. BLUE provides both unbiased and efficient estimates of the parameters, meaning that the estimates have the smallest variance among all unbiased estimators. The BLUE solution is given by:

$$\begin{aligned}\mathbf{x}_{BLUE} &= (\mathbf{H}^\top \mathbf{Q}_u^{-1} \mathbf{H})^{-1} \mathbf{H}^\top \mathbf{Q}_u^{-1} \mathbf{z} \\ \mathbb{C}[\mathbf{x}]_{BLUE} &= (\mathbf{H}^\top \mathbf{Q}_u^{-1} \mathbf{H})^{-1}\end{aligned}\tag{3.12}$$

In case the probability distribution of the measurement is known, an estimator can be obtained by maximising the likelihood function of the measurement $l(\mathbf{z}|\mathbf{x})$:

$$\mathbf{x}_{ML} = \arg \max_{\mathbf{x}} l(\mathbf{z}|\mathbf{x}),\tag{3.13}$$

Such an estimator is called the *Maximum Likelihood Estimator* (MLE). Note that if the measurement noise is Gaussian white noise, the BLUE and MLE solutions are equivalent. Similarly, the LSE is equivalent to BLUE when the measurements have similar variances. In fact, we can say that both LSE and BLUE are special cases of WLS: LSE is equivalent to WLS with unit weights, while BLUE is equivalent to WLS when the weight matrix is the inverse of the measurement covariance. It is important to note that, in a stochastic model, the error term $\mathbf{u}$ is a random variable; therefore, the resulting estimator is also a random variable when given a random input $\mathbf{z}$. However, the estimate itself is modeled as a constant deterministic variable.

3.2 Bayesian Estimation

In the non-Bayesian approach, the state variable $\mathbf{x}$ is modeled as a deterministic (fixed) value, while the measurement variable $\mathbf{z}$ is treated as a random variable. It is assumed that the observation we have is a realisation of the measurement density. This realisation is used to estimate the unknown state variable $\mathbf{x}$ and its covariance $\mathbb{C}[\mathbf{x}]$. This type of parameter estimation technique is also known as **regression**. The key distinction in the Bayesian approach is that the state variable $\mathbf{x}$ is modeled as a random variable with a *prior* density function $p(\mathbf{x})$.

Before delving into the mechanics of Bayesian estimation, it is important to grasp the role of the unknown state $\mathbf{x}$ (the parameters of interest) which changes from being a deterministic quantity to a random quantity which has a probability distribution. In **non-Bayesian** case, the likelihood in 3.13, the state $\mathbf{x}$ unknown but deterministic, and $l(\mathbf{z}|\mathbf{x})$ ¹ is PDF of the measurement $\mathbf{z}$ which is parameterised by the state $\mathbf{x}$ and the goal is to find the state $\mathbf{x}$ which maximise the measurement likelihood $l(\mathbf{z}|\mathbf{x})$. On the other hand, the likelihood function² $p(\mathbf{z}|\mathbf{x})$ in the **Bayesian** approach is a **conditional** PDF of the measurement given a realisation of random variable $\mathbf{x}$ (the state). However, both expressions ultimately provide the same value.

The first step in obtaining a Bayesian estimator is to calculate the *posterior* density³ of the state $\mathbf{x}$, given the measurement $\mathbf{z}$. The *posterior* is obtained by means of Bayes' theorem:

$$p(\mathbf{x}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{x}) p(\mathbf{x})}{\int d\mathbf{x} p(\mathbf{z}|\mathbf{x}) p(\mathbf{x})}, \quad (3.14)$$

In practice, the full posterior distribution itself may not carry much significance; it can be more useful to provide a single estimate of the state of interest, known as a **point estimates**. One could simply choose the most likely state as a point estimate, i.e., the state which maximises the posterior; the obtained point estimator, in this case, is referred to as **maximum a posteriori estimator (MAP)**. Another common choice to obtain a point estimate is to report the mean of the posterior, namely, the conditional expectation of $\mathbf{x}$ given $\mathbf{z}$.

$$\hat{\mathbf{x}} = \mathbb{E}[\mathbf{x}|\mathbf{z}] \triangleq \int_{-\infty}^{\infty} \mathbf{x} p(\mathbf{x}|\mathbf{z}) d\mathbf{x}. \quad (3.15)$$

¹Alternative notation is used to emphasise that this is not a conditional probability but rather a function of the state: $f_{\mathbf{z}}(\mathbf{z}; \mathbf{x})$.

²It is also referred to as evidence.

³It is often difficult to obtain an exact solution to the posterior, however, in most of the cases it is sufficient to calculate the expected value and the variance of the posterior density.

3.2.1 Minimum Mean Square Error Estimator

In a similar vein to the LSE estimator (3.5), an objective function or criterion can be chosen to derive an estimator for a random variable. To find a point estimate from a given density $p(\mathbf{x})$, a reasonable criterion could be *expected value of the squared estimation error*, or in other words, the **Mean Squared Error (MSE)**: $\mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top]$. This leads us to a formulation whereby, we will try to find an estimate $\hat{\mathbf{x}}$ which minimises the MSE overall possible estimate

$$\hat{\mathbf{x}} = \arg \min_{\hat{\mathbf{x}}} \mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top]. \quad (3.16)$$

This estimate can be calculated by setting the gradient (or derivative in the scalar case) of the MSE with respect to $\hat{\mathbf{x}}$ to zero

$$\begin{aligned} \nabla_{\hat{\mathbf{x}}} \mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top] &= \nabla_{\hat{\mathbf{x}}} \mathbb{E}[\mathbf{x}^\top \mathbf{x} - 2\mathbf{x}^\top \hat{\mathbf{x}} + \hat{\mathbf{x}}^\top \hat{\mathbf{x}}] \\ &= \nabla_{\hat{\mathbf{x}}} \mathbb{E}[\mathbf{x}^\top \mathbf{x}] - 2\mathbb{E}[\mathbf{x}^\top] \hat{\mathbf{x}} + \hat{\mathbf{x}}^\top \hat{\mathbf{x}} \\ &= 2\mathbb{E}[\mathbf{x}^\top] + 2\hat{\mathbf{x}}^\top = 0 \end{aligned} \quad (3.17)$$

Therefore, the point estimates $\hat{\mathbf{x}}$ of the density $p(\mathbf{x})$ which minimises the MSE, is given by

$$\hat{\mathbf{x}} = \mathbb{E}[\mathbf{x}]. \quad (3.18)$$

Although the true state $\mathbf{x}$ remains unknown, we have knowledge of its statistics. Therefore, the solution will depend on the statistics of the unknown state $\mathbf{x}$ rather than the unknown true parameter $\mathbf{x}$.

Similarly, a point estimate from the conditional density $p(\mathbf{x}|\mathbf{z})$ can be derived. For a choice of the MSE as a cost function, one can show that the **conditional expectation** $\mathbb{E}[\mathbf{x}|\mathbf{z}]$ in (3.15) is the optimal solution/estimator which minimises the MSE criteria. An estimator derived under this criteria is the so-called **Minimum Mean Square Error Estimator (MMSE)** ¹

$$\hat{\mathbf{x}}^{MMSE} = \arg \min_{\hat{\mathbf{x}}} \mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top | \mathbf{z}]. \quad (3.19)$$

¹In other branches of science, MMSE is also referred to as **Least Mean Squares Estimator (LMSE)**

By setting the gradient of (3.19) to zero with respect to $\hat{\mathbf{x}}$

$$\begin{aligned}\nabla_{\hat{\mathbf{x}}} \mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top | \mathbf{z}] &= \nabla_{\hat{\mathbf{x}}} \mathbb{E}[\mathbf{x}^\top \mathbf{x} - 2\mathbf{x}^\top \hat{\mathbf{x}} + \hat{\mathbf{x}}^\top \hat{\mathbf{x}} | \mathbf{z}] \\ &= \nabla_{\hat{\mathbf{x}}} \mathbb{E}[\mathbf{x}^\top \mathbf{x} | \mathbf{z}] - 2\mathbb{E}[\mathbf{x}^\top | \mathbf{z}] \hat{\mathbf{x}} + \hat{\mathbf{x}}^\top \hat{\mathbf{x}} \\ &= -2\mathbb{E}[\mathbf{x}^\top | \mathbf{z}] + 2\hat{\mathbf{x}}^\top = 0\end{aligned}\tag{3.20}$$

therefore, the solution to (3.19) is the conditional mean of $\mathbf{x}$

$$\hat{\mathbf{x}}_{MMSE} = \mathbb{E}[\mathbf{x} | \mathbf{z}] = \int \mathbf{x} p(\mathbf{x} | \mathbf{z}) d\mathbf{x}\tag{3.21}$$

We want to emphasise that MMSE is a Bayesian approach that calculates the mean of the posterior distribution $p(\mathbf{x} | \mathbf{z})$, which is calculated by utilising both the prior probability $p(\mathbf{x})$ and the likelihood $p(\mathbf{z} | \mathbf{x})$.

Even though the conditional mean $\mathbb{E}[\mathbf{x} | \mathbf{z}]$ is optimal for MSE criteria, it is often difficult to compute the conditional expectation $\mathbb{E}[\mathbf{x} | \mathbf{z}]$, which requires a full knowledge about the density $p(\mathbf{x} | \mathbf{z})$. To simplify the problem, a linear assumption can be made on the function the estimator can take, then derive an estimator under MSE criteria. This estimator is called **Linear Minimum Mean Square Error Estimator (LMMSE)** and is discussed in the following subsection.

3.2.2 Linear Minimum Mean Square Error Estimator

Suppose we want to estimate the state $\mathbf{x}$, given a realisation $\mathbf{z}$ of the measurement, using the MMSE method. The estimate obtained by MMSE is $\hat{\mathbf{x}} = f(\mathbf{z}) = \mathbb{E}[\mathbf{x} | \mathbf{z}]$. However, calculating $\mathbb{E}[\mathbf{x} | \mathbf{z}]$ can be challenging in practical scenarios as it requires computing the full posterior $p(\mathbf{x} | \mathbf{z})$ beforehand. Additionally, the function $f(\mathbf{z})$ may have a complex form. To address these issues, we can employ a simplified form of $f(\mathbf{z})$. Specifically, we can constrain $f(\mathbf{z})$ to a linear form, expressed as: $f(\mathbf{z}) = \beta_0 + \sum_{i=1}^n \beta_i \mathbf{z}_i$.

Definition 3.2.1 : *The LMMSE¹ estimator of a random variable $\mathbf{x}$ is a function of the form*

$$\hat{\mathbf{x}} = \beta_0 + \sum_{i=1}^n \beta_i \mathbf{z}_i\tag{3.22}$$

$$= \beta_0 + \mathbf{B}^\top \mathbf{Z},\tag{3.23}$$

¹This estimator is also known as the **best linear unbiased estimator** and as the **linear least mean squares estimator**.

which minimises the squared error $\mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top]$.

With this definition, calculating the optimal estimator became a problem of finding the linear coefficients $\beta_0, \beta_1, \dots, \beta_n$ that minimises $\mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top]$. In order to find the value of the coefficient which minimises the square error, one can derive the MSE $\mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top]$ with respect to each coefficient and set it to zero

$$\nabla_{\beta_i} \mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top] = 0. \quad (3.24)$$

This yields the expression of the **LMMSE** for the multidimensional case as:

$$\hat{\mathbf{x}} = \mathbb{E}[\mathbf{x}] + \mathbf{P}_{\mathbf{xz}} \mathbf{P}_{\mathbf{zz}}^{-1} (\mathbf{z} - \mathbb{E}[\mathbf{z}]), \quad (3.25)$$

where $\mathbf{P}_{\mathbf{zz}} = \mathbb{E}[(\mathbf{z} - \hat{\mathbf{z}})(\mathbf{z} - \hat{\mathbf{z}})^\top]$ is the covariance of the measurements, while $\mathbf{P}_{\mathbf{xz}} = \mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{z} - \hat{\mathbf{z}})^\top]$ is the cross-covariance between the state and the measurements.

The covariance of the estimator in equation (3.25) can be computed using the definition of covariance:

$$\mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top] = \mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}] + \mathbf{P}_{\mathbf{xz}} \mathbf{P}_{\mathbf{zz}}^{-1} (\mathbf{z} - \mathbb{E}[\mathbf{z}])(\mathbf{x} - \mathbb{E}[\mathbf{x}] + \mathbf{P}_{\mathbf{xz}} \mathbf{P}_{\mathbf{zz}}^{-1} (\mathbf{z} - \mathbb{E}[\mathbf{z}]))^\top] \quad (3.26)$$

after manipulating this equation, this term can be expressed as:

$$\mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^\top] = \mathbf{P}_{\mathbf{xx}} - \mathbf{P}_{\mathbf{xz}} \mathbf{P}_{\mathbf{zz}}^{-1} \mathbf{P}_{\mathbf{zx}} \quad (3.27)$$

The equations in (3.25) and (3.27) are the **fundamental equations of linear estimation**.

3.2.3 MMSE Estimator for Multivariate Gaussian

The multivariate Gaussian distribution plays an important role in sensor data fusion and state estimation. We will show the result of calculating the conditional mean of multivariate Gaussian variables, which is the MMSE estimator of multivariate Gaussian variables.

Theorem 2.2.3-1: Assume two random vector $\mathbf{x}$ and $\mathbf{z}$ are jointly Gaussian, that is,

$$p(\mathbf{x}, \mathbf{z}) = \mathcal{N}(\mathbf{y}, \bar{\mathbf{y}}, \mathbf{P}) \quad (3.28)$$

with

$$\mathbf{y} = \begin{bmatrix} \mathbf{x} \\ \mathbf{z} \end{bmatrix} \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} \mathbf{P}_{\mathbf{xx}} & \mathbf{P}_{\mathbf{xz}} \\ \mathbf{P}_{\mathbf{zx}} & \mathbf{P}_{\mathbf{zz}} \end{bmatrix} \quad (3.29)$$

Then, the conditional distribution of $\mathbf{x}$ given $\mathbf{z}$ is also Gaussian with mean

$$\mathbb{E}[\mathbf{x}|\mathbf{z}] = \mathbb{E}[\mathbf{x}] + \mathbf{P}_{\mathbf{xz}}\mathbf{P}_{\mathbf{zz}}^{-1}(\mathbf{z} - \mathbb{E}[\mathbf{z}]) \quad (3.30)$$

and covariance matrix

$$\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^\top | \mathbf{z}] = \mathbf{P}_{\mathbf{xx}} - \mathbf{P}_{\mathbf{xz}}\mathbf{P}_{\mathbf{zz}}^{-1}\mathbf{P}_{\mathbf{zx}} \quad (3.31)$$

and the marginal distributions of $\mathbf{x}$ and $\mathbf{z}$ are also Gaussian with mean vector $\mathbb{E}[\mathbf{x}]$ and $\mathbb{E}[\mathbf{z}]$ covariance matrix $\mathbf{P}_{\mathbf{xx}}$ and $\mathbf{P}_{\mathbf{zz}}$, respectively. The proof, which directly follows from the matrix inversion lemma, is straightforward but lengthy. For interested readers, the proof can be found in A.5 [28] or in [6].

3.2.4 Unbiased Estimators

We can define an estimator in any way we like; however, it is necessary to have a metric to quantify the quality of the estimator we choose.

Definition 3.2.2 *An estimator $\hat{\mathbf{x}}$ is said to be unbiased if the estimation error*

$$\tilde{\mathbf{x}} \triangleq \mathbf{x} - \hat{\mathbf{x}} \quad (3.32)$$

is a zero mean, that is,

$$\mathbb{E}[\mathbf{x} - \hat{\mathbf{x}}] = 0. \quad (3.33)$$

Example Consider a set of measurements $\mathbf{Z} = \{\mathbf{z}_1, \dots, \mathbf{z}_n\}$ of the random variable $\mathbf{x}$, where each measurement $\mathbf{z}^i \in \mathbf{Z}$ is corrupted by white Gaussian noise $\mathbf{u}^i$, such that $\mathbf{z}^i = \mathbf{x} + \mathbf{u}^i$. Let an estimator of the state $\mathbf{x}$ be

$$\hat{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{z}^i, \quad (3.34)$$

the goal is to determine whether the estimator $\hat{\mathbf{x}}$ is unbiased.

Solution: The above unbiasedness requirement for the point estimate in (3.34) is

satisfied if:

$$\mathbb{E}[\mathbf{x} - \hat{\mathbf{x}}] = 0. \quad (3.35)$$

By substituting 3.34 in 3.35 we obtain

$$\mathbb{E}[\mathbf{x} - \hat{\mathbf{x}}] = \mathbb{E}\left[\mathbf{x} - \frac{1}{N} \sum_{i=1}^N \mathbf{z}^i\right] \quad (3.36)$$

$$= \mathbb{E}\left[\mathbf{x} - \frac{1}{N} \sum_{i=1}^N (\mathbf{x} + \mathbf{u}^i)\right] \quad (3.37)$$

$$= \mathbb{E}\left[\mathbf{x} - \frac{1}{N} \sum_{i=1}^N \mathbf{x} - \frac{1}{N} \sum_{i=1}^N \mathbf{u}^i\right] \quad (3.38)$$

$$= \mathbb{E}\left[\mathbf{x} - \mathbf{x} \frac{N}{N} - \frac{1}{N} \sum_{i=1}^N \mathbf{u}^i\right] \quad (3.39)$$

$$= \mathbb{E}\left[-\frac{1}{N} \sum_{i=1}^N \mathbf{u}^i\right] \quad (3.40)$$

$$= -\frac{1}{N} \sum_{i=1}^N \mathbb{E}[\mathbf{u}^i] \quad (3.41)$$

$$= 0 \quad (3.42)$$

Thus, $\hat{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{z}^i$ is unbiased estimator of $\mathbf{x}$.

3.3 Dynamic Estimation for Target Tracking

In target tracking applications, the state of interest is dynamic and evolves over time. In this case, to solve the estimation problem, it is necessary to evolution model how the state $\mathbf{x}$ evolves over time. Just as we add a stochastic component to account for measurement noise in 3.3, a stochastic component is also incorporated into the model to account for the uncertainty associated with process noise. Typically, the evolution model done using linear equation with additive white Gaussian noise:

$$\mathbf{x}_k = \mathbf{F}_{k|k-1} \mathbf{x}_{k-1} + \mathbf{v}_k \quad (3.43)$$

where $\mathbf{F}_{k|k-1}$ is the state transition model also called evolution matrix; $\mathbf{v}_k$ is zero-mean white Gaussian process noise, with covariance, $\mathbb{E}[\mathbf{v}_k \mathbf{v}_k^\top] = \mathbf{D}_{k|k-1} : \mathbf{v}_k \sim \mathcal{N}(\mathbf{v}_k; 0, \mathbf{D}_{k|k-1})$.

Utilizing the underlying Markov assumption, we often represent the knowledge of the state with a Gaussian probability distribution of the target state at time t_k , conditioned on the previous state $\mathbf{x}_{k-1}$. This conditional distribution is referred to as the transition density, denoted as $p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \mathbf{F}_{k|k-1}\mathbf{x}_{k-1}, \mathbf{D}_{k|k-1})$. This representation allows us to predict future object states by combining the evolution model with the prior knowledge about the state by applying the Chapman-Kolmogorov equation:

$$p(\mathbf{x}_k) = \int d\mathbf{x}_{k-1} p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{z}^{k-1}), \quad (3.44)$$

where $p(\mathbf{x}_{k-1}|\mathbf{z}^{k-1}) = \mathcal{N}(\mathbf{x}_{k-1}; \mathbf{x}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})$ is the Gaussian prior density about the state. As direct result of the product formula for Gaussian A.5[28], the prediction density is also provided by a Gaussian:

$$p(\mathbf{x}_k|\mathbf{z}^{k-1}) = \int d\mathbf{x}_{k-1} p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{z}^{k-1}) \quad (3.45)$$

$$= \int d\mathbf{x}_{k-1} \mathcal{N}(\mathbf{x}_k; \mathbf{F}_{k|k-1}\mathbf{x}_{k-1}, \mathbf{D}_{k|k-1}) \mathcal{N}(\mathbf{x}_{k-1}; \mathbf{x}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}) \quad (3.46)$$

$$= \mathcal{N}(\mathbf{x}_k; \mathbf{F}_{k|k-1}\mathbf{x}_{k-1}, \mathbf{F}_{k|k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k|k-1}^\top + \mathbf{D}_{k|k-1}) \quad (3.47)$$

$$= \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1}). \quad (3.48)$$

The filtering density $p(\mathbf{x}_k|\mathbf{z}^k)$ is obtained by combining the sensor model $p(\mathbf{z}_k|\mathbf{x}_k)$ with the prediction density $p(\mathbf{x}_k|\mathbf{z}^{k-1})$:

$$p(\mathbf{x}_k|\mathbf{z}^k) = \frac{p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}^{k-1})}{\int d\mathbf{x} p(\mathbf{z}_k|\mathbf{x}) p(\mathbf{x}_k|\mathbf{z}^{k-1})}, \quad (3.49)$$

For the case where the sensor model is linear and Gaussian $p(\mathbf{z}_k|\mathbf{x}_k) = \mathcal{N}(\mathbf{z}_k; \mathbf{H}\mathbf{x}_k, \mathbf{R}_k)$, the Bayesian approach to solving this type of estimation problem can be obtained in closed form using the well-known Kalman filter update equations. Consequently, Kalman filtering can be regarded as straight-forward and direct implementation of the broader Bayesian filtering framework. As direct result of the product formula for Gaussian A.5[28], Equation (3.49) can be solved as:

$$p(\mathbf{x}_k | \mathbf{Z}^k) = \frac{p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathcal{Z}^{k-1})}{\int d\mathbf{x} p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathcal{Z}^{k-1})} \quad (3.50)$$

$$= \frac{\mathcal{N}(\mathbf{z}_k; \mathbf{H}\mathbf{x}_k, \mathbf{R}_k) \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1})}{\int d\mathbf{x} \mathcal{N}(\mathbf{z}_k; \mathbf{H}\mathbf{x}_k, \mathbf{R}_k) \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1})} \quad (3.51)$$

$$= \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k}, \mathbf{P}_{k|k}), \quad (3.52)$$

the first moment $\mathbf{x}_{k|k}$ and the second central moment $\mathbf{P}_{k|k}$ of the state $\mathbf{x}_k$ are given by:

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{W}_{k|k-1}(\mathbf{z}_k - \mathbf{H}\mathbf{x}_{k|k-1}) \quad (3.53)$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{W}_{k|k-1} \mathbf{S}_{k|k-1} \mathbf{W}_{k|k-1}^\top \quad (3.54)$$

where the Kalman Gain matrix is given by:

$$\mathbf{W}_{k|k-1} = \mathbf{P}_{k|k-1} \mathbf{H}_k^\top \mathbf{S}_{k|k-1}^{-1} \quad (3.55)$$

and the matrix $\mathbf{S}_{k|k-1}$ results from the predicted covariance matrix and the sensor model:

$$\mathbf{S}_{k|k-1} = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k \quad (3.56)$$

Note that there are equivalent formulation of the Kalman filtering equations according to various version of the product formula A.5 [28].

3.3.1 Kalman Filter vs the Fundamental Equations of Linear Estimation

Within the context of the Kalman filter, the focus is on estimating the state at time t_k , denoted as $\mathbf{x}_k$. The prior estimate of this state, stemming from the prediction step, can be described as:

$$\mathbb{E}[\mathbf{x}_k | \mathcal{Z}^{k-1}] = \mathbf{x}_{k|k-1} \quad (3.57)$$

The state $\mathbf{x}_k$ is updated based on the measurement $\mathbf{z}_k$, with prior mean — the **predicted measurement**:

$$\mathbb{E}[\mathbf{z}_k | \mathcal{Z}^{k-1}] = \mathbf{H}\mathbf{x}_{k|k-1}. \quad (3.58)$$

Also, it is required to estimate the predicted state covariance $\mathbf{P}_{k|k-1}$

$$\mathbb{E}[(\mathbf{x}_k - \mathbb{E}[\mathbf{x}_k])(\mathbf{x}_k - \mathbb{E}[\mathbf{x}_k])^\top | \mathcal{Z}^{k-1}] = \mathbf{P}_{\mathbf{xx}} = \mathbf{P}_{k|k-1} \quad (3.59)$$

The predicted measurement covariance is

$$\mathbb{E}[\mathbf{z}_k | \mathcal{Z}^{k-1}] = \mathbf{P}_{\mathbf{zz}} = \mathbf{S}_{k|k-1} \quad (3.60)$$

Now the cross covariance between the state $\mathbf{x}_k$ and the measurement $\mathbf{z}_k$ is computed as:

$$\mathbb{E}[(\mathbf{x}_k - \mathbb{E}[\mathbf{x}_k])(\mathbf{z}_k - \mathbb{E}[\mathbf{z}_k])^\top | \mathcal{Z}^k] = \mathbf{P}_{\mathbf{xz}} \quad (3.61)$$

The weighing matrix of the fundamental equation of linear estimation (static estimation) $\mathbf{P}_{\mathbf{xz}}\mathbf{P}_{\mathbf{zz}}^{-1}$ becomes the Kalman gain $\mathbf{W}_{k|k-1}$ from the update equation (3.53):

$$\mathbb{E}[(\mathbf{x}_k - \mathbb{E}[\mathbf{x}_k])(\mathbf{z}_k - \mathbb{E}[\mathbf{z}_k])^\top | \mathcal{Z}^k] \mathbb{E}[\mathbf{z}_k | \mathcal{Z}^{k-1}]^{-1} = \mathbf{P}_{\mathbf{xz}}\mathbf{P}_{\mathbf{zz}}^{-1} = \mathbf{W}_{k|k-1} \quad (3.62)$$

with this we can substitute (3.62) into the Kalman filter equations (3.53) and (3.54)

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \overbrace{\mathbf{P}_{\mathbf{xz}}\mathbf{P}_{\mathbf{zz}}^{-1}}^{\mathbf{W}_{k|k-1}}(\mathbf{z}_k - \mathbb{E}[\mathbf{z}_k]) \quad (3.63)$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{\mathbf{xx}} - \overbrace{\mathbf{P}_{\mathbf{xz}}\mathbf{P}_{\mathbf{zz}}^{-1}}^{\mathbf{W}_{k|k-1}} \overbrace{\mathbf{P}_{\mathbf{zz}}\mathbf{P}_{\mathbf{zx}}}^{\mathbf{S}_{k|k-1}} \overbrace{\mathbf{P}_{\mathbf{zz}}^{-1\top}}^{\mathbf{W}_{k|k-1}^\top} \mathbf{P}_{\mathbf{zx}} \quad (3.64)$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{\mathbf{xx}} - \mathbf{P}_{\mathbf{xz}}\mathbf{P}_{\mathbf{zz}}^{-1}\mathbf{P}_{\mathbf{zx}}, \quad (3.65)$$

The Kalman filter update equation shares similarities with static estimation, while the dynamic aspect of the system is effectively addressed by the Kalman filter prediction step.

Chapter 4

Extended Target Tracking Bayesian Framework

In this chapter, we explore the Bayesian framework for tracking a single extended target using a sequence of noisy measurements. We begin with an overview of the extended target tracking problem, followed by a detailed Bayesian formulation that includes modeling the state of the extended target, prediction, filtering, and designing the measurement likelihood associated with the target.

4.1 Overview of extended target tracking

Conventional target tracking typically assumes that received measurements are independent, with each target generating only a single measurement point per scan, effectively treating the target as a single point. The common challenges in point target tracking include missed detections, clutter, varying numbers of targets, sensor noise, and other uncertainties. However, due to advances in sensor technologies and increasing resolution, it has become essential for tracking systems with high-resolution sensors to account for the spatial extent of the target. These systems must also recognise extended targets as distinct entities, initiate extended target tracks, and effectively maintain these tracks over time.

Applications that benefit from this approach include object tracking using cameras, vehicle tracking for driving assistance, and maritime tracking of boats and ships (see Fig. 4.1). The spatial extent of a target should be considered an additional degree of freedom that characterises an extended target and must be estimated simultaneously with its kinematic properties. Therefore, the target extension becomes a part of the

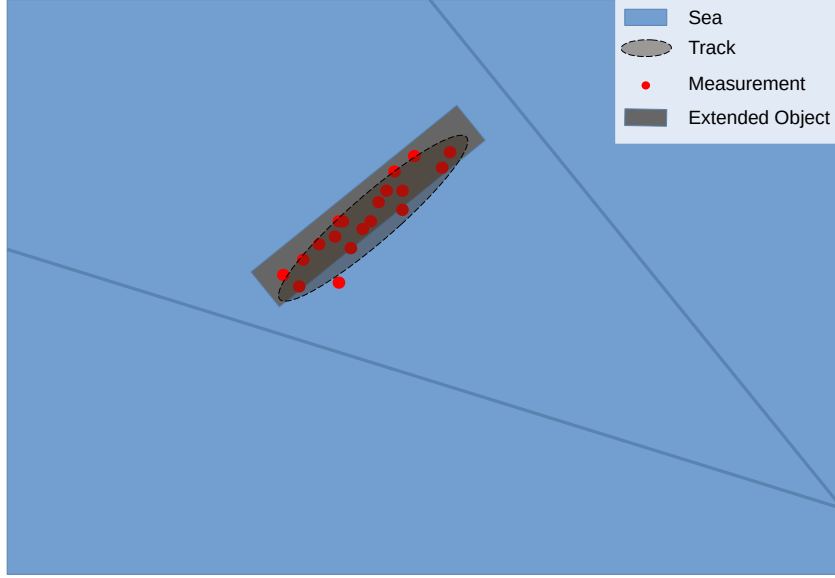


Fig. 4.1 Depicts a marine scenario featuring simulated radar detections from a marine surveillance radar. It also shows an extended target track with a shape moulded as an ellipse.

object's state. Given this, we can define the extended target tracking problem as follows:

Definition 4.1.1 *An extended target is a target whose rigid body¹ spans multiple resolution cells of a sensor, potentially generating multiple measurements in each sensor scan.*

Definition 4.1.2 *Extended target tracking (ETT) aims to address the challenge of simultaneously estimating both the kinematics and the extent/shape² of a target based on a sequence of noisy measurements.*

Over the past decade, this challenge has captured the attention of researchers, resulting in several proposed solutions aimed at overcoming it. Additionally, multiple extended target tracking (METT) poses challenges in data association, as many association events must be considered. The METT problem is beyond the scope of this work, as our focus is on estimating both the kinematic parameters and the shape of objects. For further details on ETT and METT please see [19, 18].

¹We add the term "rigid body" here to distinguish it from group targets, as some literature treats groups and extended targets as the same.

²In the work, the terms "target shape" and "target extent" are used interchangeably.

It's important to note here that a related problem arises when tracking a group of closely-spaced objects. Due to the resolution capabilities of the sensor and the sensor-to-target geometry, two or more targets within the group can be indistinguishable. In such cases, it's crucial to treat the group as a single entity and estimate both its shape and kinematics. However, the main difference between extended target tracking and group target tracking is that the shape of the group is not rigid. Additionally, the group can form and dissolve, requiring additional considerations to address these points. Examples of group target tracks include aircraft formations, ground-moving convoys, and drone swarms.

4.1.1 Bayesian formalism

The philosophy of the Bayesian framework is to compute the probability of a random variable of interest based on prior knowledge and measurable evidence. With this, we can formulate a general Bayesian tracking algorithm for extended targets by computing the conditional probability $p(\mathbf{x}_k|\mathcal{Z}^k)$ at each time t_k given sensor data $\mathcal{Z}^k$. The extended target state vector $\mathbf{x}_k$ should include both kinematic and shape variables. The state typically includes the following:

- Kinematic state: The kinematic parameters of the targets such as position, velocity, acceleration, heading, and turn-rate.
- Extent state: The parameters determining the shape and orientation of the object.

An example of a 2D extended target state is:

$$\mathbf{x}_k = [x_k, y_k, x'_k, y'_k, l_k, w_k, \theta_k]^\top \quad (4.1)$$

where x_k, y_k represent the 2D Cartesian position, x'_k, y'_k are the Cartesian velocity components and l_k, w_k, θ_k denote the length, width, and orientation of the target, respectively. Extended target tracking involves iteratively calculating the conditional density $p(\mathbf{x}_k|\mathcal{Z}^k)$, which comprises two primary steps: prediction and filtering.

Prediction

In view of the discussion in Sect. 3.3, the target evolution model describes how the target state evolves over time. For a point target, the evolution model usually involves modelling the evolution of the position and other kinematic parameters. However, in

ETT, the model should also involve the evolution of the extent over time. Similar to conventional point target tracking, prediction is achieved by combining the evolution model $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ with the prior state density $p(\mathbf{x}_{k-1})$ using Chapman-Komogorove equation (3.44). The evolution model for the kinematic state employs standard point target tracking dynamic models, such as constant velocity (CV), constant acceleration (CA), and constant turn-rate (CT); for a comprehensive overview, refer to [37].

The extended target, as defined in Definition 4.1.1, maintains a rigid body, meaning its size and shape remain constant over time. However, the sensor may not perceive the full shape due to self-occlusion, leading to variations of the shape appearance depending on sensor-to-target geometry and target orientation. Depending on the chosen extent model, a suitable transition density must be designed, although this can introduce significant complexity. In many applications like tracking cars bicycles or ships, a simple symmetric geometric shape such as an ellipse or rectangle is sufficient. Also, it is typical to assume that the orientation of the target is aligned with the heading therefore, the orientation doesn't need to be explicitly added as a new degree of freedom to the model. With such an assumption the prediction of the transition model of the shape can be as simple as an identity model with suitable process noise. For more than one spatial model refer to [19].

Filtering

After the prediction, there is a filtering stage where the set of the reflected measurement from the target $\mathcal{Z}_k = \{\mathbf{z}^i\}_{i=0}^{n_z}$ at time t_k needs to be processed. Specifically, within this filtering step, a sensor-specific and extended target state-specific likelihood function $p(\mathcal{Z}_k|\mathbf{x}_k)$ is defined and combined with the predicted density $p(\mathbf{x}_k|\mathcal{Z}^{k-1})$ by utilizing Bayes' formula (3.50).

Measurement model

To compute the likelihood of a measurement $p(\mathcal{Z}_k|\mathbf{x}_k)$, we must establish a model that specifies the location of the measurement source on the target surface. This model is contingent upon factors such as the target state vector, the sensor type, and the shape used to represent the extent. The selection of the shape model needs to be made carefully to match the required level of shape detail in a given application. Estimating and tracking complex shapes adds a lot of complexity to the tracking algorithm, which can be avoided in many applications where basic geometric shapes such as an ellipse, a line, or a rectangle are sufficient to describe the object. However, in other applications where making assumptions about the shape is difficult, computationally complex algo-

rithms with a general model might be needed to track and describe random shapes. Fig. 4.2 provides an example illustrating the differences between LiDAR and radar reflection points in an elliptical model of the shape. Moreover, sensor noise influences

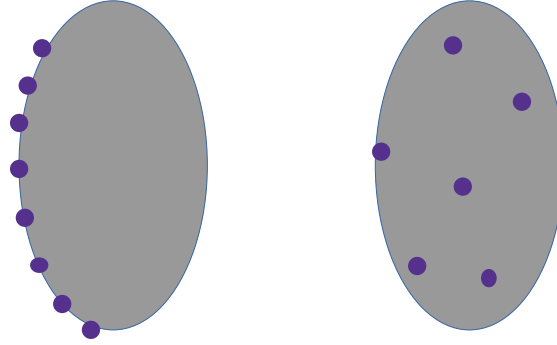


Fig. 4.2 Illustrate possible measurement sources of an elliptical model for the shape of a LiDAR sensor located on the left side of the object in the left figure, and a radar located on the left side of the object in the right figure.

the detections, necessitating the incorporation of these properties into the measurement modeling process. Typically, sensor noise is addressed by modeling uncertainties using probabilistic tools. The model should contain both the number of detections and their spatial distribution on the target extent. The most common ways to model ET measurement likelihood are:

- **Spatial model:** In this framework, measurements are assumed to be distributed across the rigid body of the target according to a probability distribution. An earlier and commonly used method, known as the Random Matrix approach [25, 14], models the extent of the target using an inverse Wishart density, resulting in the representation of the target shape as an ellipse. Alternatively, another approach treats each detection $\mathbf{z}$ as a noisy measurement of a source $\mathbf{y}$ positioned somewhere on the target [7]. In [21], the authors proposed a model for tracking cars using LiDAR sensors, where the source density $p(\mathbf{y}|\mathbf{x})$ is assumed to be uniform along the sides that are visible to the sensor. A Gaussian density was used for the noise $p(\mathbf{z}|\mathbf{y})$, thus, the measurement likelihood is the marginalisation

of the reflection point $\mathbf{y}$ of the object:

$$p(\mathcal{Z}_k|\mathbf{x}) = \prod_{\mathbf{z} \in \mathcal{Z}_k} \int p(\mathbf{z}|\mathbf{y})p(\mathbf{y}|\mathbf{x})d\mathbf{y} \quad (4.2)$$

- Set of measurement sources: This model assumes that there are d independent measurement sources (reflection points) located on the target body [22, 23]. This approach necessitates solving the data association problem for each measurement source with the received measurements before conducting the update.

4.2 Extended Target Tracking: Spatial Elliptical Model for Laser Measurements

In this chapter, we introduce a method outlined in our publication [4] for tracking extended targets using LiDAR measurements. Here, the sources of the measurements are distributed along the contour of the target’s extent. We employ an Extended Kalman Filter (EKF) to estimate both the target’s kinematics and shape simultaneously. The target’s shape is represented by an ellipse, and the measurement sources are uniformly distributed along the ellipse’s contour. To approximate the measurement likelihood, we utilise a Gaussian mixture. The EKF update incorporates a moment-matching technique to approximate the mixture density. We evaluate the performance of the proposed approach using data from a LiDAR simulator to track a target with a rectangular shape.

4.2.1 Introduction

With high-resolution sensors like LiDAR, overlooking the spatial extent of the target is an oversimplification of the issue. According to Definition 4.1.2, the challenge of extended target tracking lies in the complexity of simultaneously estimating the kinematics and shape of both targets based on multiple scans. In each scan, the target produces a set of measurements reflected from spatially distributed measurement sources on the target surface, as illustrated in Fig. 4.3.

Sophisticated shape models in extended target tracking can accommodate targets with unknown or arbitrary shapes, where there is no explicit model of the target’s geometrical shape. One pioneering approach in this context is the Random Hypersurface Model (RHM) [8, 13, 9]. RHM represents the shape with a radius function that maps an angle and a shape parameter vector defined by a Fourier series to a radius, representing

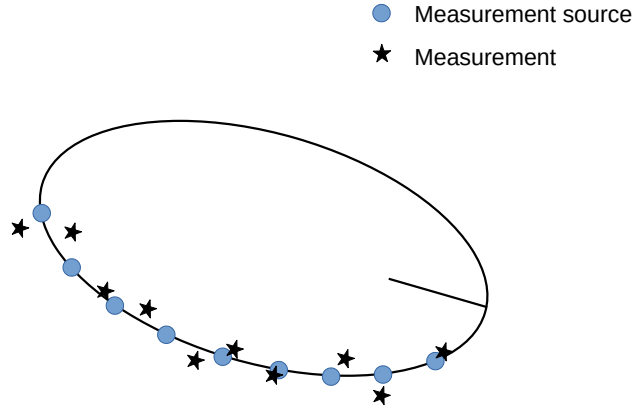


Fig. 4.3 Illustrates a set of noisy measurements resulting from measurement sources lying on the contour of an elliptical target.

the contour point at a given angle. This simplifies the problem to curve fitting. In [47], a novel approach utilising a radius function has been developed in the form of a Gaussian processes model instead of Fourier coefficients. Another common approach in the literature uses a B-spline to represent the shape [51, 24]. However, for tracking vehicle applications, a less sophisticated shape model that assumes a basic geometrical shape of the target extent, such as an ellipse or a rectangle, is sufficient. The most common approach in this context is the Random Matrix approach [25, 14], which assumes an elliptical shape of the target, where the prior of the target extent is modeled as an Inverse Wishart distribution. In [20], rectangular and elliptical shapes are used to model the target extent, where the measurement source is spatially distributed around a predicted part of the target shape. The kinematics and shape update was performed using an extended Kalman filter together with Gaussian mixture probability hypothesis density (GM-PHD). However, there was no explicit mathematical function for the measurement model. Thus, the authors computed the Jacobian for the measurement model numerically instead of deriving it analytically.

In this approach, the target shape is assumed to be an ellipse with measurement sources uniformly distributed around the contour of the ellipse. Therefore, it is customised for use with a LiDAR sensor. The measurement model is derived from the standard parametric function of the ellipse, enabling the computation of predicted measurements and their corresponding innovation covariances within the extended

Kalman filter framework. The extended target measurement likelihood is represented by a Gaussian mixture density. Consequently, standard nonlinear estimation methods such as EKF could be utilised to estimate the extended target state vector.

4.2.2 Target Extent Model

This section illustrates the single extended target state vector and the measurement model used in this work.

Extended Target State Vector

At time t_k , the state vector of the extended target $\mathbf{x}_k \in \mathbb{R}^6$ holds both kinematic and shape variables

$$\mathbf{x}_k = [x_k, y_k, x'_k, y'_k, l_{1,k}, l_{2,k}]^\top \quad (4.3)$$

where

- x, y, x', y' are the kinematic variables that represent the position and the velocity of the target (and potentially additional variables),
- $l_{1,k} \in \mathbb{R}$ and $l_{2,k} \in \mathbb{R}$ represent the ellipse semi-axis, thus their absolute value represents the target semi-length and width.

Measurement Model

At discrete instants of time t_k , the extended target reflects n_z noisy measurements $\mathcal{Z}_k = \{\mathbf{z}^i\}_{i=0}^{n_z}$, originating from a set of measurement sources $\mathcal{Y} = \{\mathbf{y}^j\}_{j=1}^m$ spatially distributed on the target surface. Considering a LiDAR sensor type, $\mathcal{Y}$ is assumed to be uniformly distributed on the contour of the extended target. Each measurement $\mathbf{z}_k^i$ stems from one of the m sources along the contour and is corrupted by white Gaussian noise $\mathbf{u}_k^i$ characterised by a measurement error covariance matrix $\mathbf{R}_k^i$ as shown in Fig. 4.4.

Therefore, a measurement $\mathbf{z}_k^i$ resulting from a measurement source $\mathbf{y}_k^i$ can be obtained by following the measurement equation:

$$\mathbf{z}_k^i = \mathbf{y}_k^i + \mathbf{u}_k^i \quad (4.4)$$

To calculate the measurement model $p(\mathcal{Z}_k|\mathbf{x}_k)$, we first need to find a model that describes the locations of measurement sources on the target surface, depending on

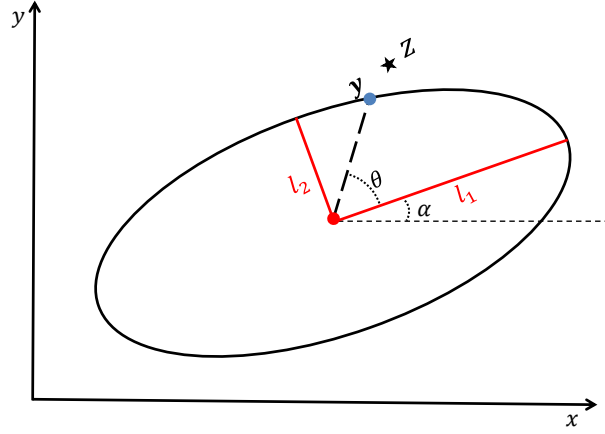


Fig. 4.4 Illustrates an elliptical target rotated by an angle α and a noisy measurement $\mathbf{z}$ originating from a particular measurement source $\mathbf{y}$ lying on the contour of the ellipse at an angle θ .

the target state and shape variables. In this work, the target shape is assumed to be an ellipse, which allows us to derive an explicit measurement equation.

Following the standard ellipse equation in a Cartesian coordinate system, any point $\mathbf{p} \in \mathbb{R}^2$ which lies on the line of an axis-aligned ellipse can be obtained by

$$\mathbf{p} = [l_1 \cos(\theta), l_2 \sin(\theta)]^\top, \quad (4.5)$$

where l_1 is the ellipse semi-major axis, l_2 is the ellipse semi-minor axis and θ is the angle of the point $\mathbf{p}$ with respect to the ellipse major axis Fig. 4.4. Writing (4.5) in matrix formulation gives us

$$\begin{bmatrix} p_x \\ p_y \end{bmatrix} = \begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \quad (4.6)$$

The rotation and the translation of the extended target need to be incorporated in the measurement equation as well as estimated. The target translation $\mathbf{c} \in \mathbb{R}^2$ from the coordinate origin coincides with the estimated target position variables of the state vector

$$\mathbf{c} = \begin{bmatrix} \mathbf{I}_2 & \mathbf{0}_4 \end{bmatrix} \mathbf{x}_k \quad (4.7)$$

Also, it is natural to assume that the orientation of the target is aligned with the direction of its velocity vector which can be realised directly from the kinematic variables in the state vector which relate to the velocity. By incorporating the translation and rotation, one can rewrite equation (4.6) as:

$$\begin{aligned}\mathbf{y} &= \mathbf{c} + \text{Rot}(\alpha) \begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \\ &= g(\mathbf{x}, \theta),\end{aligned}\tag{4.8}$$

where $\text{Rot}(\alpha)$ is the rotation matrix corresponding to the angle α of the velocity vector. Then, the measurement equation can be obtained by incorporating the measurement noise:

$$\mathbf{z} = \mathbf{c} + \text{Rot}(\alpha) \begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} + \mathbf{u}\tag{4.9}$$

Therefore, the measurement likelihood becomes:

$$p(\mathbf{z}|\mathbf{x}, \theta) = \mathcal{N}(\mathbf{z}; g(\mathbf{x}, \theta), \mathbf{R})\tag{4.10}$$

The angle θ in (4.8) is an input in a bounded interval $\theta \in [0, 2\pi]$, although its exact value is unknown for a particular measurement source $\mathbf{y}$. However, as the value of θ is bounded, the measurement likelihood $p(\mathbf{z}|\mathbf{x})$ can be modeled as a uniform distribution along the ellipse line. Thus, $p(\mathbf{z}|\mathbf{x})$ is obtained by marginalising the possible values of the angle θ within the interval, that is:

$$p(\mathbf{z}|\mathbf{x}) = \int_0^{2\pi} d\theta p(\mathbf{z}|\mathbf{x}, \theta) p(\theta|\mathbf{x})\tag{4.11}$$

We may approximate the integral in (4.11) by a Riemann sum, which results in m components Gaussian mixture with equal weights $w^j = \frac{1}{m}$

$$p(\mathbf{z}|\mathbf{x}) \approx \sum_{j=1}^m w^j \mathcal{N}(\mathbf{z}; g(\mathbf{x}, \theta^j), \mathbf{R})\tag{4.12}$$

The number of Gaussian mixture components m reflects the number of measurement sources and is a parameter choice that corresponds to the number of sample points placed within the interval $[0, 2\pi]$. Fig. 4.5 shows a heat map of the measurement likelihood function resulting from the Gaussian mixture approximation.

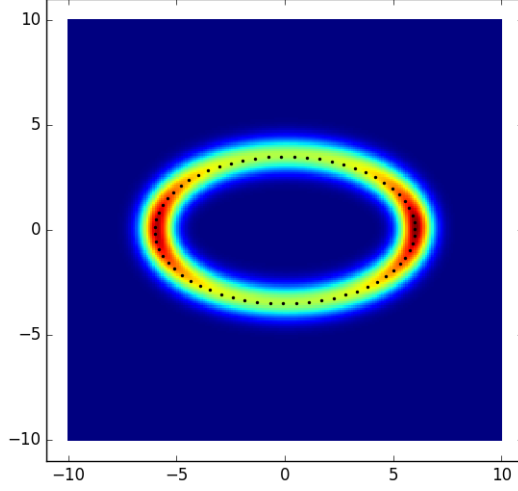


Fig. 4.5 This figure shows a heat map representing the probability density of the Gaussian mixture with the number of components corresponding to the measurement sources shown by the black points

4.2.3 Measurement Update

The update equation for the state conditioned on the measurement $\mathcal{Z}^k$ up to the current time stamp t_k is represented by the density function $p(\mathbf{x}_k|\mathcal{Z}^k)$. According to Bayes' rule:

$$p(\mathbf{x}_k|\mathcal{Z}^k) = p(\mathbf{x}_k|\mathcal{Z}_k, \mathcal{Z}^{k-1}), \quad (4.13)$$

$$= \frac{p(\mathcal{Z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathcal{Z}^{k-1})}{\int d\mathbf{x}_k p(\mathcal{Z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathcal{Z}^{k-1})} \quad (4.14)$$

$$\stackrel{\text{iid}}{=} \frac{\prod_{i=1}^{n_z} p(\mathbf{z}_k^i|\mathbf{x}_k) p(\mathbf{x}_k|\mathcal{Z}^{k-1})}{\int d\mathbf{x}_k p(\mathcal{Z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathcal{Z}^{k-1})} \quad (4.15)$$

$$(4.16)$$

Under the assumption that the measurement likelihood is a Gaussian, the Bayesian formulation leads us directly to the Kalman filter formulation

$$p(\mathbf{x}_k|\mathcal{Z}^k) = \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k}, \mathbf{P}_{k|k}) \quad (4.17)$$

$$= \frac{\prod_{i=1}^{n_z} \sum_{j=1}^m w^j \mathcal{N}(\mathbf{z}_k^i; g(\mathbf{x}_k, \theta^j), R) \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1})}{\int d\mathbf{x}_k \prod_{i=1}^{n_z} \mathcal{N}(\mathbf{z}_k^i; g(\mathbf{x}_k, \theta^j), \mathbf{R}) \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1})} \quad (4.18)$$

However, as it can be seen in (4.9) the measurement equation is nonlinear, therefore it is required to linearise the measurement equation so that the Kalman filter assumption

can hold. A Taylor expansion up to the second order provides a linear approximation of the measurement equation concerning $\mathbf{x}_k$ at a given point $\mathbf{x}_{k|k-1}$ as follows:

$$g(\mathbf{x}_{k|k}, \theta) = g(\mathbf{x}_{k|k-1}, \theta) + \nabla_{\mathbf{x}} g(\mathbf{x}_{k|k-1}, \theta)(\mathbf{x}_{k|k} - \mathbf{x}_{k|k-1}) \quad (4.19)$$

where $\nabla_{\mathbf{x}} g(\mathbf{x}_{k|k-1}, \theta)$ is the Jacobian matrix containing the partial derivatives of $\nabla_{\mathbf{x}} g(\mathbf{x}_{k|k-1}, \theta)$ which will be denoted in the rest of the section as $\mathbf{J}$ and it is formulated as:

$$\mathbf{J} = \frac{\partial g(\mathbf{x}_{k|k-1}, \theta)}{\partial \mathbf{x}} \quad (4.20)$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & \cos \alpha \cos \theta & -\sin \alpha \sin \theta \\ 0 & 1 & 0 & 0 & \sin \alpha \cos \theta & \cos \alpha \sin \theta \end{bmatrix} \quad (4.21)$$

Note that the partial derivatives in respect to the velocity components x' and y' depend on the estimated dimensions, velocity and angle θ . Although the partial derivatives in respect to x' and y' are reduced to zero to avoid double counting of information and calculate a Jacobian independent of the estimated dimensions.

One of the advantages of approximating the measurement likelihood by a Gaussian mixture lies in the fact that the Kalman filter update can easily be generalised to incorporate the Gaussian mixture. Incorporating the Gaussian mixture in (4.18) and writing the update equation only up to its proportional term, yields the following expression:

$$p(\mathbf{x}_k | \mathbf{z}_k) \propto \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1}) \sum_{j=1}^m w^j \mathcal{N}(\mathbf{z}_k; g(\mathbf{x}_k, \theta^j), \mathbf{R}) \quad (4.22)$$

Therefore, the updated density $p(\mathbf{x}_k | \mathbf{z}_k)$ can be expressed as a sum of weighted Gaussians

$$p(\mathbf{x}_k | \mathbf{z}_k) = \sum_{j=1}^m w^j \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k}^j, \mathbf{P}_{k|k}^j) \quad (4.23)$$

for a Gaussian mixture of this form, *moment matching* can be applied directly to approximate the probability density $p(\mathbf{x}_k | \mathbf{z}_k)$

$$p(\mathbf{x}_k | \mathbf{z}_k) \approx \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k}, \mathbf{P}_{k|k}) \quad (4.24)$$

where $\mathbf{x}_{k|k}$ and $\mathbf{P}_{k|k}$ are given by:

$$\mathbf{x}_{k|k} = \sum_{j=1}^m w^j \mathbf{x}_{k|k}^j \quad (4.25)$$

$$\mathbf{P}_{k|k} = \sum_{j=1}^m w^j (\mathbf{P}_{k|k}^j + (\mathbf{x}_{k|k}^j - \mathbf{x}_{k|k})(\mathbf{x}_{k|k}^j - \mathbf{x}_{k|k})^\top) \quad (4.26)$$

From the EKF formulation, the quantities $\mathbf{x}_{k|k}^j$ and $\mathbf{P}_{k|k}^j$ are to be calculated as follows:

$$\mathbf{x}_{k|k}^j = \mathbf{x}_{k|k-1} + \mathbf{W}_k^j \mathbf{v}_k^j \quad (4.27)$$

$$\mathbf{P}_{k|k}^j = \mathbf{P}_{k|k-1} - \mathbf{W}_k^j \mathbf{S}_k^j \mathbf{W}_k^{j\top} \quad (4.28)$$

with:

$$\mathbf{v}_k^j = \mathbf{z}_k - g(\mathbf{x}_{k|k-1}, \theta^j) \quad (4.29)$$

$$\mathbf{W}_k^j = \mathbf{P}_{k|k-1} \mathbf{J}_k^{j\top} \mathbf{S}_{k|k-1}^{j-1} \quad (4.30)$$

$$\mathbf{S}_{k|k-1}^j = \mathbf{J}_k^j \mathbf{P}_{k|k-1} \mathbf{J}_k^{j\top} + \mathbf{R} \quad (4.31)$$

and the weight is calculated as:

$$w^j = \frac{\mathcal{N}(\mathbf{z}; g(\mathbf{x}, \theta^j), \mathbf{S}_{k|k-1}^j)}{\sum_{j=1}^m \mathcal{N}(\mathbf{z}; g(\mathbf{x}, \theta^j), \mathbf{S}_{k|k-1}^j)} \quad (4.32)$$

4.2.4 Prediction

In this work a linear motion model is used, so standard Kalman filter equations are used to calculate the predicted state $\mathbf{x}_{k|k-1}$ and covariance $\mathbf{P}_{k|k-1}$ as

$$\mathbf{x}_{k|k-1} = \mathbf{F} \mathbf{x}_{k-1|k-1} \quad (4.33)$$

$$\mathbf{P}_{k|k-1} = \mathbf{F} \mathbf{P}_{k-1|k-1} \mathbf{F}^\top + \mathbf{D} \quad (4.34)$$

where $\mathbf{F}$ is the matrix that defines the deterministic motion model, which is given by:

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 1 & 0 & \Delta t & 0 & 0 \\ & \mathbf{0}_{4 \times 2} & & \mathbf{I}_{4 \times 4} & & \end{bmatrix} \quad (4.35)$$

and $\mathbf{D}$ matrix characterises the motion covariance.

4.2.5 Results

In this section, the proposed approach is tested and evaluated using a simulation framework which is designed to model Ibeo LiDAR characteristics; such as, the number of reflected measurements, measurement noise, and other characteristics. At each simulation frame, a stationary LiDAR mounted at position $(0, 0)$ produces a number of measurements reflected from the contour of the target. The ground truth shape of the simulated extended target used in this work is a vehicle with length $= 5.6m$ and width $= 1.9m$, thus the shape resembles a rectangle with rounded corners. Fig. 4.7 illustrates the result of the discussed approach in a simulation scenario which combines linear and curved trajectories covered by the sensor surveillance area. Also, in this section the proposed approach which we will be referred to as EKF approach is compared to the Random Matrix-Feldmann approach [14]. Both the Random Matrix approach and the proposed approach assume an elliptic shape for the target. Except that, in the Random Matrix approach the measurements are assumed to be normally distributed on the target surface rather than the target boundary as in the proposed approach.

The evaluation of the shape and the kinematic estimation is carried out independently. An *Intersection over Union* (IoU) metric which measures the accuracy of an estimated rectangle is utilised in this work to evaluate the shapes. In order to apply an IoU metric, the area of the target ground truth bounding box A and the estimated area of the track bounding box $\hat{A}$ has to be calculated. IoU is calculated as follows:

$$\text{IoU} = \frac{A \cap \hat{A}}{A \cup \hat{A}} \in [0, 1] \quad (4.36)$$

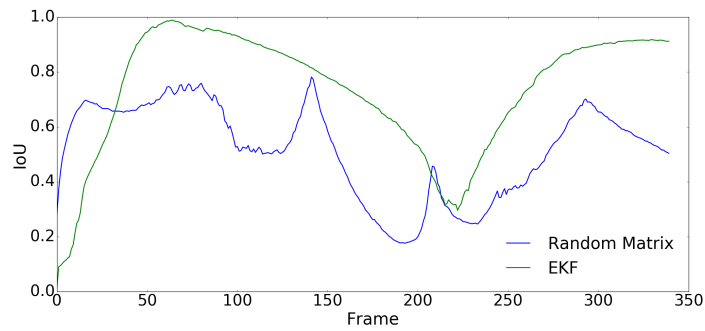
Thus, IoU computes the ratio of the intersection over the union of the two bounding boxes. The ratio will be close to 1 in case the areas A and $\hat{A}$ are heavily overlapping, or close to 0 otherwise. In this work, the shape is estimated as an ellipse, however,

extracting a bounding box out of the ellipse is straightforward where the length and the width of the bounding box are set to the same value of the estimated ellipse major and minor axes respectively.

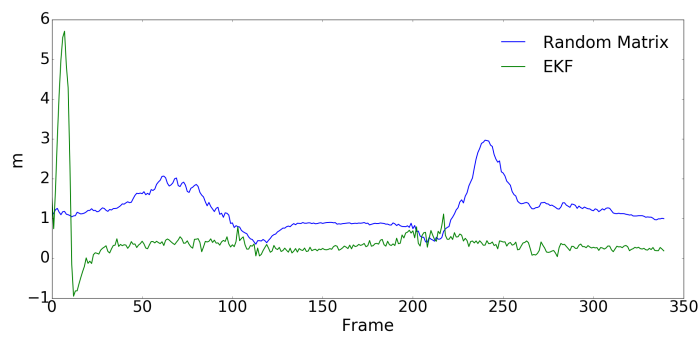
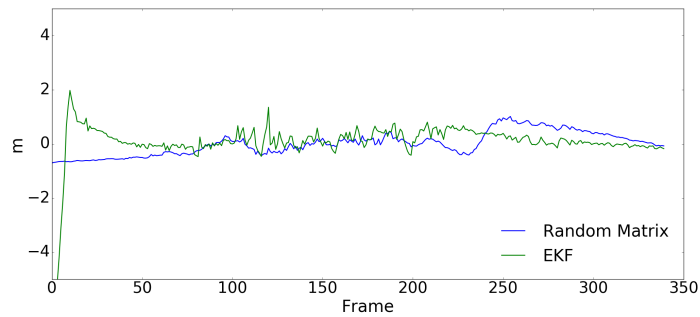
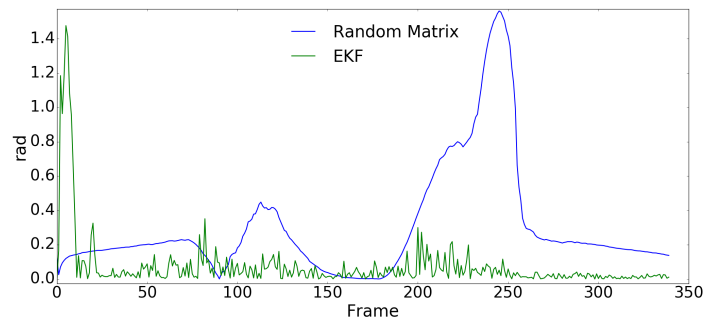
The target starts and continues in a linear motion for the first quarter of the frames during this period the IoU ratio illustrated in Fig. 4.6(a), shows that EKF approach has a steep increase in the accuracy of the estimated shape. After reaching more than 95% accuracy, the IoU ratio starts to drop constantly for almost two thirds of the frames. That is due to the fact that the number of the received measurement points from the target decreases as the target distancing from the sensors. Another reason for the drop of the curve results from the target nonlinear motion where it is clearly seen a significant drop in IoU ratio at the frames where the target had change in its orientation. In the remaining frames, and as the target starts approaching the sensor in a linear motion, the accuracy of the estimated shape starts to recover. The graphs in Fig. 4.6(b) and (c) show about a 0.5m estimation error in average of the x and y positions. It is believed that it is due to the fact that the estimator fits an ellipse only to the two parts which were seen by the sensor that caused the estimator to be biased to the seen part of the target. This estimation error is calculated by taking the difference between the centres of the estimated target and the target ground truth. The estimation error of the orientation illustrated in Fig. 4.6(d) represents the difference between the ground truth orientation of the target and direction of the velocity vector which, in our work, represents the estimated orientation of the target. As the velocity is not directly correlated with the shape, a very small estimation error of the orientation is presented.

As expected the EKF approach shows better results overall in comparison to the Random Matrix in the different tests Fig. 4.6. That can be explained by the Random Matrix assuming that the measurements are generated from the centre of the target rather than its contour. Thus the target centre is estimated with a high bias toward the centre of the received measurements. Also, the Random Matrix approach appears to suffer largely from the change of orientation, which is a result of the Random Matrix not estimating the orientation explicitly. In other words, the Random Matrix does not distinguish between a change in orientation and a change in size.

To better illustrate the result of the proposed approach, Fig. 4.7 shows the EKF estimated track by the red ellipse compared to the Random Matrix estimate represented by the blue ellipse and to the ground truth shown by the black rectangle. As it can be seen the overall performance is good and the behaviour of the ellipse fitting of the measurement points correctly reflects the proposed measurement model in (4.9).



(a) IoU ratio

(b) x position error(c) y position error

(d) Orientation error

Fig. 4.6 The figure illustrates the evaluation results of the proposed approach as shown by the blue curve compared to the Random Matrix approach shown by the green curve.

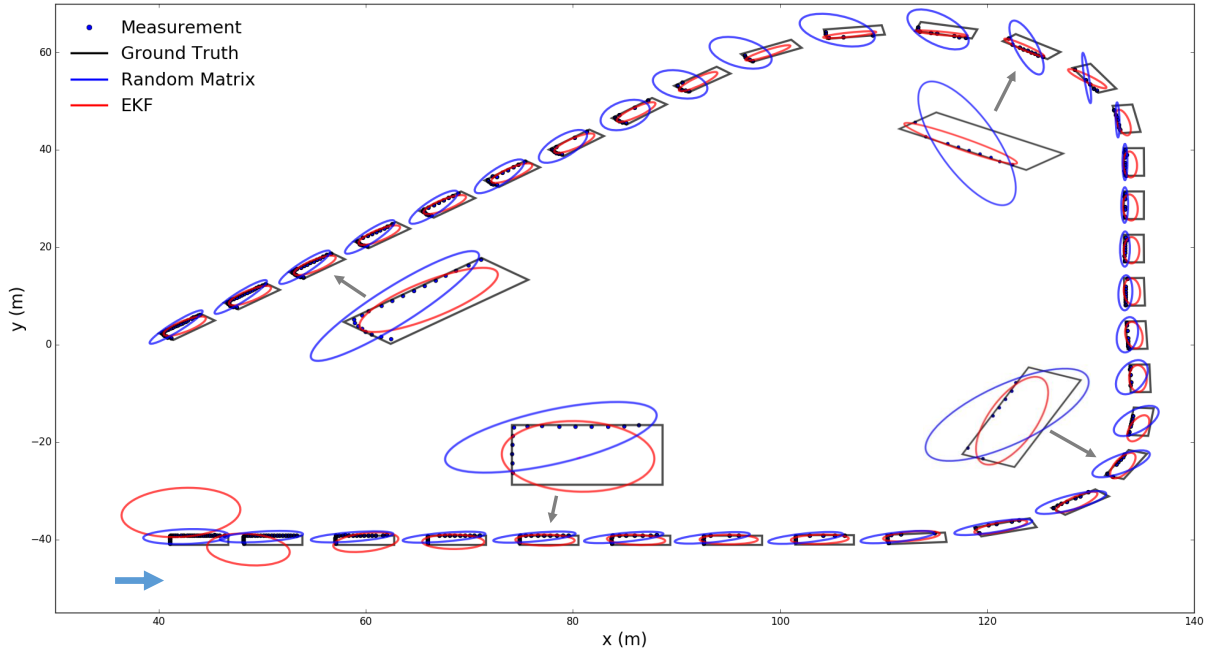


Fig. 4.7 In this simulation result, the single extended target ground truth is represented by the black rectangle and the dark blue points show the measurements reflected from the target to a LiDAR sensor located at coordinate $(0,0)$. The estimated target shape by proposed approach is represented by the red ellipse, while the Random Matrix-Feldmann approach is represented by the blue ellipse.

4.2.6 Conclusion

This section presented an approach which is capable of estimating both kinematic and shape variables of an extended target simultaneously. The proposed approach presented an explicit mathematical function of the measurement model derived directly from the standard geometrical function of the assumed shape. The predicted measurement in the form of a Gaussian mixture was computed and incorporated in EKF framework. The presented results show that the approach can be applied to track target with rectangular shape with good accuracy. Even though the assumption of elliptical extents of the target seems restrictive to track an object with elliptical or rectangular extents, it is still a very good fit for certain tracking application, such as tracking cars, boats, etc. One of the limitations of this approach, when only one side of the target is seen up to the current time stamp, the estimated ellipse will have almost zero value for one of its axes as the ellipse will try to fit a measurements distribution which resembles a line. The future plan is to investigate a solution for this limitation.

4.3 Extended Target Tracking: With Positive Support Density Filtering

In many data fusion and state estimation applications, the parameter of interest inherently assumes positive values, such as in distance estimation or counting instances of certain items. In such cases, optimal data fusion requires modeling the system state as a positive random variable, with its probability density function confined to the positive real axis. Conventionally, in stochastic estimation problems, where we seek to estimate an unknown state based on measurements from noisy sensory devices like radar or LiDAR. A common assumption is that the random error in the measurements follows a white Gaussian noise model. Consequently, the unknown state is often described as a random variable with a Gaussian density. However, in certain scenarios, such as when the unknown state is known to be strictly positive (e.g., estimating size or dimension), classical approaches based on Gaussian densities may fail, potentially yielding negative values.

This chapter discusses modeling quantities that are constrained to positive values with a distribution that has positive support. The uni-variate Gamma distribution has positive support and represents the maximum entropy distribution for such variables. To facilitate Bayesian recursion with Gamma density, an approximate moment-matching approach is proposed. Additionally, we explore a special case of the multivariate Gamma distribution, namely the Wishart distribution. We demonstrate how it can be used in a recursive Bayesian framework to estimate a multivariate random variable, thereby modeling the unknown state as a Wishart-distributed random matrix. This modeling choice ensures that the probability densities of the random variables remain restricted to positive real values. Then the feasibility of the proposed Bayesian filters are demonstrated within the framework of Extended Target Tracking.

4.3.1 Introduction

In many data fusion and estimation applications, the state of interest only takes positive values. For example, it might be the goal to estimate a distance or size. In such a case the classical approaches based on Gaussian densities might produce erroneous results, in particular whenever the variance of the likelihood is rather large compared to the mean. Optimal data fusion then should model the system state as a positive random variable, which has a probability density function whose support is restricted to the positive real axis, that is, $\int_0^\infty p(x) dx = 1$. The vast majority of model-based methods

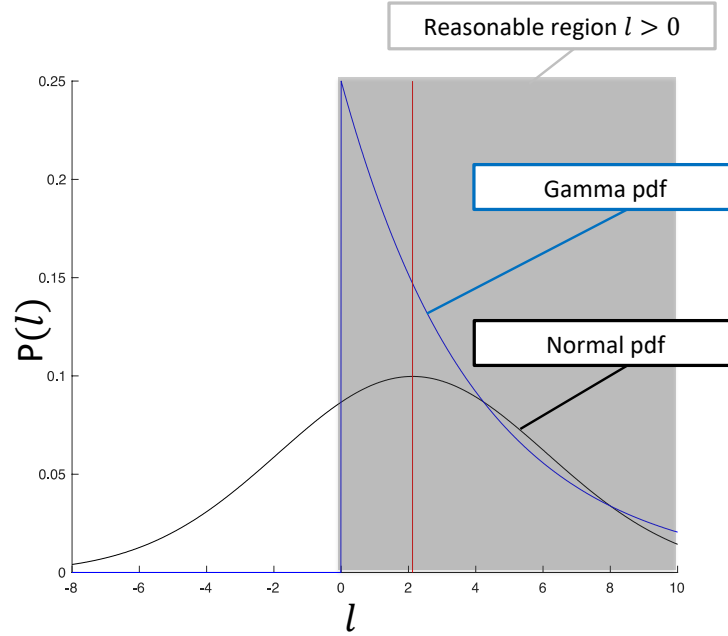


Fig. 4.8 Exemplary distributions of a parameter $d = 2.1$ with a standard deviation $\sigma = 4$. The Gaussian density (black) has a significant part of its probability mass within the negative real axis whereas the Gamma distribution is limited to the positive real axis.

rely on the Bayes formula, as it provides the means to integrate information of a data time series from multiple potentially heterogeneous and spatially distributed sensors and also from contextual background knowledge [28]. The Kalman filter has become the ‘working horse’ of Bayesian tracking and state estimation over the years for various reasons:

- Its Gaussian assumption is suitable for many applications.
- It’s computationally fast and easy to implement.
- Its limitations often can be overcome by various extensions.

However, the assumption of a Gaussian distributed state can yield an insufficient representation of the actual knowledge of a state of interest if the latter is known to be positive, that is, its distribution is limited to the positive real axis. A demonstrative example is given in Fig. 4.8,

where a parameter l is given with its true value 2.1. Assume we have given an unbiased estimate with a standard deviation $\sigma = 4$. From the figure, it becomes

obvious that the normal distribution $\mathcal{N}(l; \mu = 2.1, \sigma = 4)$ has a significant part of its probability mass on the negative real axis which contradicts our knowledge that d can only take positive numbers.

There is a wide range of applications, which put up the challenge of positive parameter estimation:

- Counting: Applications with counting instances include crowd estimation, epidemiological models, or industrial quantity or volume computation.
- Distances: Measuring distances is a ubiquitous sensor application based on radar, LiDAR, ultra sound, or stereoscopic imagery.
- Intensities: Intensity estimation appears for instance in signal processing using acoustic or electromagnetic signals.
- Errors: To estimate the absolute error, standard deviation, or the variance, positive densities are required.

This behaviour is not limited only to the theoretical sphere, but it also appears clearly when applied. A direct application of positive parameter estimation appears in ETT problems [19], which require estimating both the target's kinematics state and shape parameters. The vast majority of ETT models in the literature assume a Gaussian distribution of the target's extent state which is known to be positive. These models suffer from the drawback of the Gaussian assumption illustrated in the example mentioned earlier – where the target extent is occasionally estimated as negative. One of the pioneer approaches in this context is the Random Hypersurface Model (RHM)[8, 13, 9]. RHM represents the shape by a radius function that maps an angle and a shape parameter vector (defined by a Fourier series) to a radius which represents the contour point at a given angle. A negative estimate of the radius length has been observed and reported in [44], where the authors suggested a solution by imposing inequality constraints on the parameters that are known to be positive.

Another common model [47], which also used a radius function, has been developed in the form of Gaussian Processes (GP) instead. The authors of the GP model reported that the simulation results indicate that the GP model is more robust and the issue of negative extent estimation occurs considerably less frequently than in RHM. The same drawback also exists in [4], where a measurement model under an elliptical shape for a LiDAR sensor assumption was suggested. This model assumes a Gaussian distribution of the ellipse axes. However, when a negative estimate of the axis occurs, the filter estimates the opposite angle's sign of the corresponding axis, thereby eliminating

its effect. Another common approach in literature uses a B-spline under Gaussian assumption to track a *star-convex* shape presented in [51, 24]. However, the Gaussian assumption compromises the robustness of this approach and others mentioned before.

A probability density with positive support has been used in the context of Radar to track the target extent. The most common approach in this sphere is the random matrix approach [25, 14], where the prior of the target extent is modeled as an Inverse Wishart distribution. In the following years, a few authors suggested some adaptation to this approach to track the target's extent using LiDAR measurements [41, 38]. This model describes the target extent in terms of a symmetric, positive definite (SPD) matrix, thus, the dimension can take only positive values. However, the target extent representation is restricted to an elliptical shape.

It's worth noting that the Poisson density function is widely used for positive random variables. For example, the Probability Density Hypothesis (PHD) Filter [32] or the Intensity Filter [43] for multi-target tracking relies on the Poisson distribution per position in the state space. However, it's evident that solely propagating and computing the first moment is often insufficient for accurately representing knowledge, especially when the parameter is small but its variance is large, or vice versa. The cardinalized extension, known as the CPHD filter [31], computes the joint density of the set event of spatially distributed instants and their cardinality. A balance between the simplicity of the PHD and the versatility of the CPHD filter is achieved through a formulation of the distribution of Panjer classes [12, 39]. It's also possible to impose constraints on a Bayes filter [50] and on Gaussians [42] to restrict the prior density function to the positive real axis. However, this comes at the expense of higher computational costs.

Based on our publications [16, 1, 3], this chapter presents an approximate solution, based on moment matching, for the prediction-update recursion of a Gamma random variable, referred to as the "Gamma filter". Additionally, we demonstrate that the exact formula is biased and derive an unbiased version, which is then applied in a simulative example. Furthermore, we introduce an extension of the Gamma filter, utilising a special case of the multivariate Gamma distribution known as the Wishart distribution [5]. We derive a Bayes update rule for the Wishart density and employ it to track the shape of an extended target in a closed-form solution. In this work, the target extent model assumes an unknown or arbitrary shape and is tailored for use with a LiDAR sensor by assuming the measurement sources to be uniformly distributed around the contour of the shape.

4.3.2 Uni-variate Gamma

Problem Formulation

The Gamma distribution of a positive random variable x is defined as the density function for parameters k^1 and θ with:

$$\mathcal{G}(x; k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} \quad (4.37)$$

where Γ is the Gamma function, k is the ‘shape’ and θ is the ‘scale’ parameter. Its mean and variance, respectively, are given by

$$\mathbb{E}[x] = k \theta \quad (4.38)$$

$$\mathbb{C}[x] = k \theta^2 \quad (4.39)$$

Some examples (for $k = 1$, $k = 2$, and $k = 4$) for a fixed scale $\theta = 1$ are shown in Fig. 4.9, where one can see that the shape depends on the ratio of its parameters. The Gamma distribution can be matched for a random variable x with a given expectation and variance

$$\mathbb{E}[x] = \mu \quad (4.40)$$

$$\mathbb{C}[x] = \sigma^2 \quad (4.41)$$

by solving for k and θ . One directly obtains

$$k = \frac{\mu^2}{\sigma^2} \quad (4.42)$$

$$\theta = \frac{\sigma^2}{\mu} \quad (4.43)$$

In this work, it is assumed that the sensor model for an observation z_k at time instant t_k of a positive parameter x_k is given in terms of a Gamma distribution:

$$\ell(z_k; x_k) = \mathcal{G}(x_k; k_z, \theta_z) \quad (4.44)$$

¹Readers need to pay attention to the different fonts used for k and k , as they refer to different values - the former representing the time step and the latter denoting the shape parameter of the Gamma distribution.

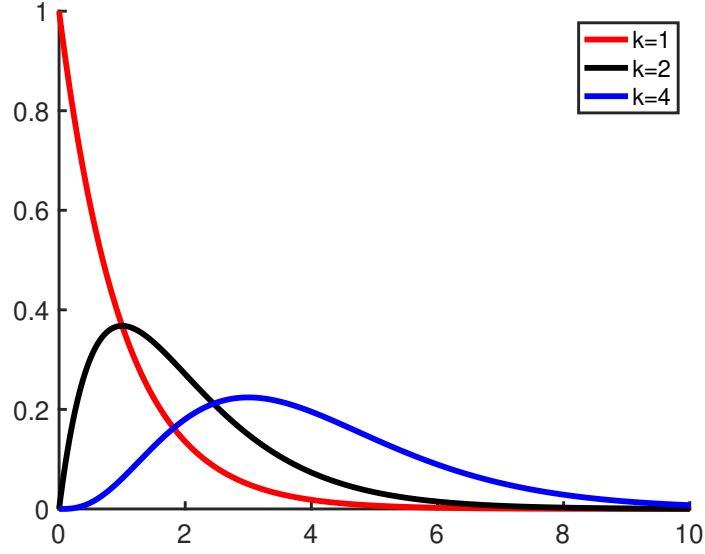


Fig. 4.9 Example plots of the Gamma distributions for different parameters, where $\theta = 1$ is fixed.

The conjugate prior for the Gamma distribution is a Gamma distribution again, therefore, the Bayesian recursion is initialised with a Gamma distribution with initial parameters k_0 and θ_0 :

$$p(x_0) = \mathcal{G}(x_0; k_0, \theta_0) \quad (4.45)$$

It is assumed that the time evolution can be described in terms of a stochastic process such that it can be integrated for the prediction step, in which the prior for time step k is computed:

$$p(x_{k-1} | \mathcal{Z}^{k-1}) \xrightarrow{\text{evolution model}} p(x_k | \mathcal{Z}^{k-1}), \quad (4.46)$$

Then, in the filtering step, the computation of the posterior is applied by means of the Bayes theorem, in which the current sensor observation z_t and its sensor model is incorporated:

$$p(x_k | \mathcal{Z}^{k-1}) \xrightarrow[\text{sensor model}]{\text{current sensor data}} p(x_k | \mathcal{Z}^k). \quad (4.47)$$

Prediction – Filtering Recursion for the Gamma Filter

For a closed Prediction–Filtering recursion, these two steps have to be solved such that the resulting density function stays within the family of Gamma distributions.

Prediction During the time evolution from time $k - 1$ to time k the parameter may undergo some deterministic changes, given by the function $f(x)$ and some additive white noise u_t with variance σ_u^2 with

$$x_t = f(x_{k-1}) + u_k. \quad (4.48)$$

Since the Gamma distribution is uniquely defined by its first and second moment, moment matching may be applied to obtain the density function with the desired properties. By using (4.40) and (4.41), the previous posterior parameters $x_{k-1|k-1} := \mathbb{E}[x_{k-1}|\mathcal{Z}^{k-1}]$ and $\sigma_{k-1|k-1}^2 := \mathbb{C}[x_{k-1}|\mathcal{Z}^{k-1}]$ are obtained directly. Analogously to the Kalman filter, the first and second moment of the predicted variable is given by

$$x_{t|t-1} = f(x_{t-1|t-1}), \quad (4.49)$$

$$\sigma_{t|t-1}^2 = f(x_{t-1|t-1})^2 \sigma_{k-1|k-1}^2 + \sigma_u^2. \quad (4.50)$$

A back transformation to a Gamma distribution yields the desired parameters for the prior distribution:

$$p(x_k|\mathcal{Z}^{k-1}) = \mathcal{G}(x_k; k_{k|k-1}, \theta_{k|k-1}), \quad (4.51)$$

$$k_{k|k-1} = \frac{x_{k|k-1}^2}{\sigma_{k|k-1}^2} \quad (4.52)$$

$$\theta_{k|k-1} = \frac{\sigma_{k|k-1}^2}{x_{k|k-1}}. \quad (4.53)$$

Filtering In the filtering step, the prior is combined with the likelihood for the current observation z_t by means of the Bayes theorem. It is assumed that the likelihood is given in terms of a Gamma distribution with parameters k_z and θ_z :

$$\ell(z_k; x_k) = \mathcal{G}(x_k; k_z, \theta_z). \quad (4.54)$$

Applying Bayes theorem yields

$$p(x_k | \mathcal{Z}^k) = \frac{p(x_k | \mathcal{Z}^{k-1}) \cdot \ell(z_k; x_k)}{\int dx_k p(x_k | \mathcal{Z}^{k-1}) \cdot \ell(z_k; x_k)}. \quad (4.55)$$

By multiplying the prior and likelihood, one obtains

$$p(x_k | \mathcal{Z}^k) \propto x_k^{k_{k|k-1} - k_z - 2} e^{-\frac{x_k}{\theta_{k|k-1}} - \frac{x_k}{\theta_z}} \quad (4.56)$$

Therefore, it follows that the posterior is a Gamma density function again with parameters $k_{k|k}$ and $\theta_{k|k}$, where

$$k_{k|k} = k_{k|k-1} + k_z - 1 \quad (4.57)$$

$$\theta_{k|k} = \frac{\theta_{k|k-1} \theta_z}{\theta_{k|k-1} + \theta_z}. \quad (4.58)$$

Since the Gamma distribution is normalised, it is not required to solve the integral in the denominator in (4.55).

Unbiased Gamma Update

Despite the fact that the update derivation in (4.57) and (4.58) are exact, it should be noted that application of the Bayes rule to Gamma distributions is not unbiased. This can be seen by the following example. Consider a density $p(x) = \mathcal{G}(x; k, \theta)$ with $\mathbb{E}[x] = k \theta$. Then the n -time fusion of p is given by

$$p(x|n) := \frac{p(x)^n}{\int dx p(x)^n}. \quad (4.59)$$

According to the update equations, $p(x|n)$ is a Gamma distribution with parameters

$$k_n = nk - n \quad (4.60)$$

$$\theta_n = \theta/n. \quad (4.61)$$

As a consequence, one obtains $\mathbb{E}[x|n] = k \theta - \theta < k \theta$. Therefore, this effect can be undesirable in particular, if the likelihood is obtained via moment matching in (4.42)

and (4.43). However, an unbiased update rule can easily be found with

$$\mathbf{k}_{k|k}^* = \mathbf{k}_{k|k-1} + \mathbf{k}_z \quad (4.62)$$

$$\theta_{k|k}^* = \frac{\theta_{k|k-1} \cdot \theta_z}{\theta_{k|k-1} + \theta_z}. \quad (4.63)$$

The effect of this rule shall be illustrated by the following example. Let the parameters of the prior and the likelihood be given by their mean and standard deviation σ as summarised in Table 4.1. From the results in the last two lines, one can see that the mean of the exact update is smaller than the corresponding parameters of the prior and likelihood. This is different when using the unbiased rule from above. The corresponding densities are presented in Figure 4.10, where one can see the point-wise multiplication with normalisation for the exact rule, while the mass is shifted towards the likelihood in the unbiased case.

	Mean	σ	k	θ
===== Prior	1.5	2.0	1.13	1.33
Likelihood	2.5	2.0	1.56	1.6
Posterior	1.22	0.94	1.69	0.73
Posterior unbiased*	1.95	1.19	2.69	0.73

Table 4.1 Example for the Bayes update for a given prior and likelihood using the exact and the unbiased rules.

Simulation

In this section, we illustrate the results obtained from testing the proposed approach on automotive applications, namely extended target tracking. The extended target tracking problem can be summarised as the problem of tracking the targets' kinematics and extent simultaneously based on multiple scans. In this test, it is assumed that the targets' kinematics and the target's extent are independent. Also the target shape is assumed to be a rectangle, logically, the dimensions of the rectangle are positive quantities. Thus, the elements of the extent state vector are modeled as Gamma distributed random variables. Following this model, the target extent is tracked using

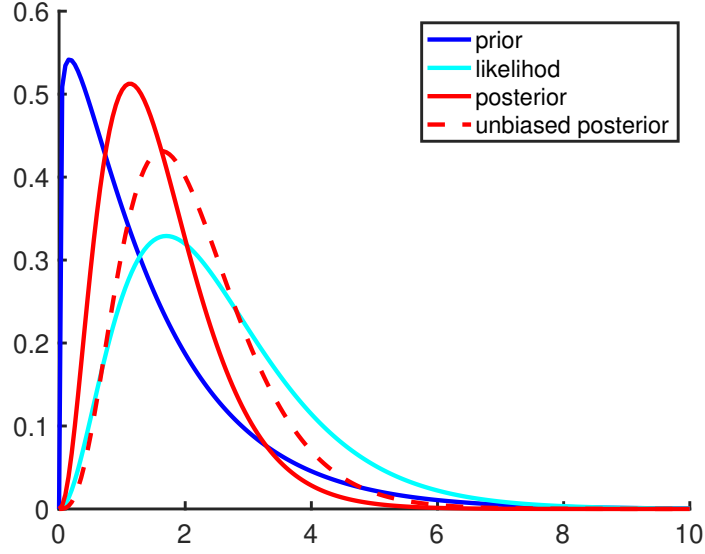


Fig. 4.10 Plots of the Gamma distributions for the exact and unbiased update rules with the parameters from Table 4.1

the proposed Gamma filter approach, while a Kalman filter with a non linear evolution model is utilised to estimate the target kinematic vector.

We used for this test a simulation framework designed to model the physical characteristics of Ibeo automotive LiDAR; such as, the number of reflected measurements, measurement noise, field of view, and other characteristics. The LiDAR sensors are mounted on a moving vehicle (ego vehicle). At each frame a point cloud reflected from measurement sources in the surrounding environment is received by the sensors. Then the point cloud is clustered to separate the measurements which belong the target to be updated.

The simulation used resembles an urban scenario, which combines the linear and curved trajectories of the targets as well as the ego vehicle. The tracking results of the linear and curved trajectories are shown by Figure. 4.11 (a)(b) and Figure. 4.11 (c)(d), respectively.

An *Intersection over Union* (IoU) metric which measures the accuracy of an estimated rectangle is utilised in this work to evaluate the shapes. In order to apply an IoU metric, the area of the target ground truth bounding box A and the estimated area of the track bounding box $\hat{A}$ has to be calculated. IoU is calculated as follows:

$$\text{IoU} = \frac{A \cap \hat{A}}{A \cup \hat{A}} \in [0, 1] \quad (4.64)$$

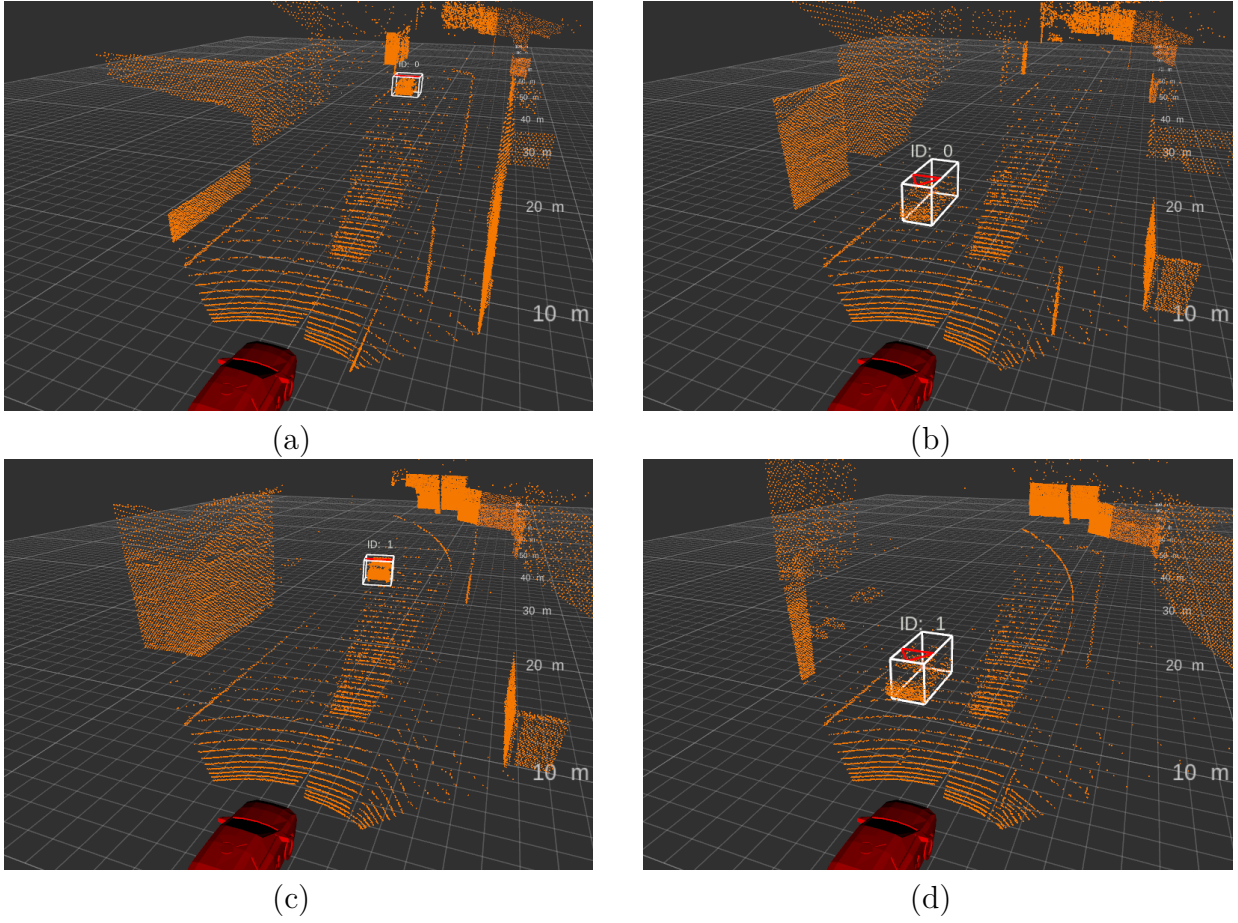


Fig. 4.11 Different snapshots are taken from a simulation of Ibeo automotive LiDAR in an urban scenario: The white frame box represents the extent of the obtained tracks using the proposed Gamma filter approach. The orange points show the LiDAR point cloud relative to the red vehicle (the ego vehicle).

Thus, IoU computes the ratio of the intersection over the union of the two bounding boxes. The ratio will be close to 1 in case the areas A and $\hat{A}$ are heavily overlapping, or close to 0 otherwise. The other metric used is the *Root Mean Square Error* (RMSE), which calculates the square root of the average squared error of N run as:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=0}^N (l - \hat{l})^2} \quad (4.65)$$

where l and $\hat{l}$ are the true and the estimated dimension, respectively.

In figure 4.18 it is shown that the IoU ratio of the target with a straight trajectory (track id zero) has a better performance compared to the target with a curved trajectory (track id one). Also, the RMSE metric shows that the width has a much smaller

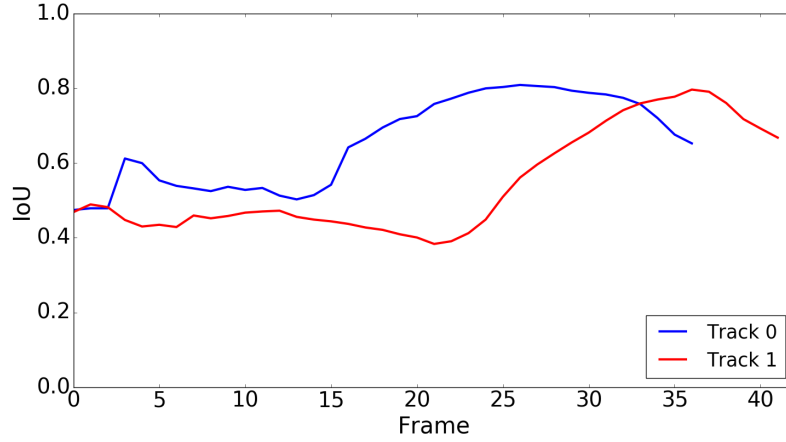


Fig. 4.12 The *Intersection over Union* results of the shape tracking using the proposed Gamma filter as shown by the blue curve which corresponds to track with id zero in figure 4.11 (a) and (b). Also by the red curve which corresponds to track id one in figure 4.11 (c) and (d).

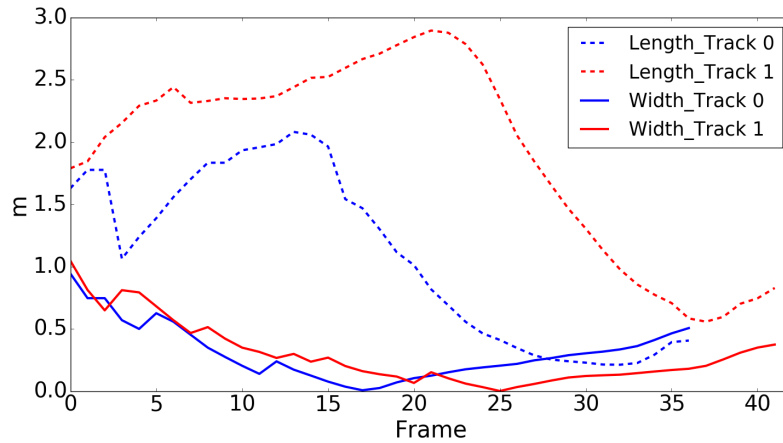


Fig. 4.13 This figure illustrates the *Root Mean Square Error* results of the estimated target's length and width. Where the track's number in the figure's legend corresponds to the track's id in figure 4.11.

error compared to the length of the target. These errors are quite explainable as at the beginning both targets are far from the sensor, and due to the sensor-to-target geometry, the length of the target is partially visible to the sensor. However, as the target starts approaching the sensor the target's length becomes more visible to the sensor, consequently, the length estimate error drops sharply. For the same reason, the IoU ratio improves. The slight drop in the performance in both metrics in the last few frames is due to the fact that the targets became outside the field of view.

4.3.3 Multivariate Gamma and Wishart Distribution

We discussed the uni-variate Gamma Bayesian filter in the previous section. In this section, we will discuss generalising the Gamma Bayesian filter to a multivariate distribution. There are different extensions of multivariate gamma distributions; not all are defined by density functions[33, 5]. However, according to [46], the generalised multivariate gamma distribution density function can be defined as:

$$\mathbf{MG}(\mathbf{X}; \alpha, \beta, \mathbf{C}) = \frac{|\mathbf{X}|^{\alpha - \frac{d+1}{2}}}{\beta^{d\alpha} \Gamma_p(\alpha) \mathbf{C}^\alpha} e^{-\frac{1}{\beta} \text{tr}(\mathbf{C}^{-1} \mathbf{X})} \quad (4.66)$$

Where $\mathbf{X}$ is a positive-definite $d \times d$ matrix, $\alpha > \frac{d-1}{2}$ shape parameter, $\beta > 0$ scale parameter, and Γ_p is the multivariate gamma function. With the specific choices of $\alpha = \frac{n}{2}$ and $\beta = 2$, the resulting PDF is known as the Wishart density. Therefore, the Wishart distribution is a special case of the multivariate Gamma distribution.

Definition 4.3.1 Suppose $\mathbf{z}$ is a $d \times n$ matrix, each j_{th} column $\mathbf{z}^j = [z^{1,j}, z^{2,j}, \dots, z^{d,j}]^\top$ is an independent realisation from a Multivariate Gaussian distribution with zero mean and dimensions d : $\mathbf{z} \sim \mathcal{N}(\mathbf{z}; 0; \mathbf{C})$. The matrix $\mathbf{z}$ is thus written as:

$$\mathbf{z} = \begin{bmatrix} z^{1,1} & \dots & z^{1,n} \\ \vdots & \ddots & \vdots \\ z^{d,1} & \dots & z^{d,n} \end{bmatrix}. \quad (4.67)$$

The joint distribution of the elements of the matrix

$$\mathbf{X} = \mathbf{z} \mathbf{z}^\top = \sum_{j=1}^n \mathbf{z}^j \mathbf{z}^{j\top} \quad (4.68)$$

is Wishart distributed with expected value $\mathbb{E}[\mathbf{X}] = n\mathbf{C}$, and is denoted by $\mathcal{W}(\mathbf{X}; n, \mathbf{C})$. The probability density function of $\mathbf{X}$ is given by:

$$\mathcal{W}(\mathbf{X}, n, \mathbf{C}) = \frac{|\mathbf{X}|^{\frac{n-d-1}{2}}}{2^{\frac{nd}{2}} \Gamma_d(\frac{n}{2}) |\mathbf{C}|^{\frac{n}{2}}} e^{-\frac{1}{2} \text{tr}(\mathbf{C}^{-1} \mathbf{X})} \quad (4.69)$$

The density $\mathcal{W}$ exists in case $n \geq d$. Where n is a positive integer representing the degrees of freedom or the sample size, $\mathbf{C}$ is a scale matrix that represents the covariance of the measurement vector $\mathbf{z}$ and $\Gamma_d(\cdot)$ is the multivariate gamma function [35].

Also, the Wishart distribution is often considered a multivariate generalisation of the chi-squared distribution (a special case of the gamma distribution).

Problem Formulation

We consider the problem of estimating a state $\mathbf{x} = [x^1, x^2, \dots, x^d]^\top$ based on a time series of noisy measurements $\mathcal{Z}^k$, where the state $\mathbf{x}$ is known to take only positive value $\mathbf{x} \in \mathbb{R}_{\geq 0}^d$. The measurement $z^i \in \mathbb{R}^1$ of an element x^i of the state $\mathbf{x}$ results from the following measurement equation

$$z_k^i = f(x^i) + u_k^i. \quad (4.70)$$

which maps x^i and an additive noise term u_k^i to the measurement z_k^i . By stacking the measurements in a vector $\mathbf{z}_k \in \mathbb{R}^d$, that is, $\mathbf{z}_k = [z_k^1, z_k^2, \dots, z_k^d]^\top$, the joint distribution of $\mathbf{z}_k$ is assumed to be Gaussian with zero mean (in reference to a given centre) and covariance matrix $\mathbf{C}$, that is, $\mathbf{z}_k \sim \mathcal{N}(\mathbf{z}; 0, \mathbf{C})$. Let's consider at each time instant t_k a number $n_z \geq d$ of an independent realisations drawn from the d-variate Gaussian $\mathcal{N}(\mathbf{z}; 0, \mathbf{C})$ are received and form the measurement matrix $\mathcal{Z}_k \in \mathbb{R}^{d \times n_z}$, such as:

$$\begin{aligned} \mathcal{Z}_k &= \begin{bmatrix} \mathbf{z}_k^1 & \cdots & \mathbf{z}_k^{n_z} \end{bmatrix} \\ &= \begin{bmatrix} z_k^{1,1} & \cdots & z_k^{1,n_z} \\ \vdots & \ddots & \vdots \\ z_k^{d,1} & \cdots & z_k^{d,n_z} \end{bmatrix}. \end{aligned}$$

Thus, $z_k^{i,j}$, $i = 1, \dots, d$, $j = 1, \dots, n_z$ represents the j^{th} measurement of the i^{th} state variable at time k . To illustrate this in an example, let's consider a state $\mathbf{x} = [x^1, x^2]^\top$ where it's elements represent the dimensions r^1 and r^2 of a box centred at the origin $(0, 0)$, Fig. 4.14. A photogrammetric technique is utilised to measure the position of the edges' points/pixels with respect to the centre of the box. In this example, the quantities r^1 and r^2 which we wish to estimate, are strictly positive, while the measurements could possibly take positive or negative values, based on which side the detected edge lies. By taking the square of the measurement matrix $\mathbf{X} = \mathbf{z}\mathbf{z}^\top = \sum_{j=1}^{n_z} \mathbf{z}^j \mathbf{z}^{j\top}$, it is guaranteed that the estimate is strictly positive. In fact, according to definition 4.3.1, the matrix $\mathbf{X} = \mathbf{z}\mathbf{z}^\top$ is Wishart distributed. In view of this, the goal is to recursively estimate a symmetric positive definite matrix $\mathbf{X}$, which follows a Wishart distribution and contains the information about the target extent given all measurements up to time k . Moreover, the Wishart distribution is a special case of the multivariate Gamma distribution and it is worth mentioning that its inverse form is widely used to track targets' extent and estimate a spread matrix [10, 14, 45].

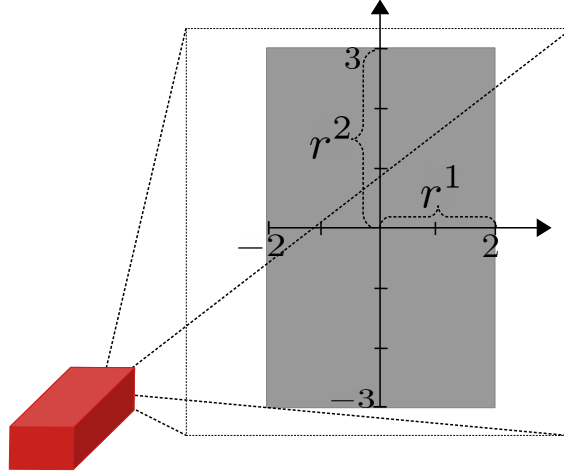


Fig. 4.14 Photogrammetric technique is used to take a spatial measurement from photographs to measure the pixels' positions related to the selected features, in this case, the edges of the object.

Wishart Distribution and Bayesian Tracking

Bayesian framework provides an iterative updating scheme to estimate stochastic variable $\mathbf{X}$ by calculating conditional probability density functions $p(\mathbf{X}_k|\mathcal{Z}^k)$ that represent all available knowledge on $\mathbf{X}$ up to time k . The related density iteration process is referred to as prediction and filtering. It is assumed that the time evolution can be described in terms of a stochastic process such that it can be integrated for the prediction step, in which the prior for time step k is computed:

$$p(\mathbf{X}_{k-1}|\mathcal{Z}^{k-1}) \xrightarrow{\text{evolution model}} p(\mathbf{X}_k|\mathcal{Z}^{k-1}), \quad (4.71)$$

Then, in the filtering step, the computation of the posterior is applied by means of the Bayes theorem, in which the current sensor observation $\mathcal{Z}_k$ and its sensor model is incorporated:

$$p(\mathbf{X}_{k-1}|\mathcal{Z}^{k-1}) \xrightarrow[\text{sensor model}]{\text{current sensor data}} p(\mathbf{X}_k|\mathcal{Z}^k). \quad (4.72)$$

Recursive Wishart Filter

In the filtering step, the likelihood of the current observation $p(\mathcal{Z}_k|\mathbf{X}_k)$ is combined with the prior density by means of the Bayes theorem:

$$p(\mathbf{X}_k|\mathcal{Z}^k) = \frac{p(\mathbf{X}_k|\mathcal{Z}^{k-1}) \cdot p(\mathcal{Z}_k|\mathbf{X}_k)}{\int d\mathbf{X}_k p(\mathbf{X}_k|\mathcal{Z}^{k-1}) \cdot p(\mathcal{Z}_k|\mathbf{X}_k)}. \quad (4.73)$$

To determine the posterior's density $p(\mathbf{X}_k|\mathcal{Z}^k)$, the densities of the measurement likelihood $p(\mathcal{Z}_k|\mathbf{X}_k)$ and the prior $p(\mathbf{X}_k|\mathcal{Z}^{k-1})$ need to be known. In our case, the sensor model exists, observations are made and $\mathcal{Z}_k$ is fixed. Thus the distribution $p(\mathcal{Z}_k|\mathbf{X}_k)$ is now considered as a likelihood function of $\mathbf{X}_k$: $\ell(\mathbf{X}_k|\mathcal{Z}_k)$ and is given in terms of a Wishart distribution characterized by the parameters n_z and $\mathbf{C}_z$ ¹:

$$\begin{aligned} p(\mathcal{Z}_k|\mathbf{X}_k) &= \ell(\mathbf{X}_k|\mathcal{Z}_k) \\ &= \mathcal{W}(\mathbf{X}_k; n_z, \mathbf{C}_z), \end{aligned} \quad (4.74)$$

where n_z is the number of measurements, and $\mathbf{C}_z$ is calculated as:

$$\mathbf{C}_z = \frac{1}{n_z} \sum_{j=1}^{n_z} \mathbf{z}_k^j \mathbf{z}_k^{j\top}. \quad (4.75)$$

The prior density is assumed to have the same family of distribution as the measurement likelihood, that is, the prior is Wishart distributed with parameters $n_{k|k-1}$ and $\mathbf{C}_{k|k-1}$. Thus, the Bayesian recursion is initialized with a Wishart distributed prior with initial parameters $\mathbf{C}_0$ and n_0 :

$$p(\mathbf{X}_0) = \mathcal{W}(\mathbf{X}_0; n_{0|0}, \mathbf{C}_{0|0}). \quad (4.76)$$

To calculate the Bayesian posterior density, first the numerator in (4.73) is solved by multiplying the prior and the likelihood densities, that yields:

$$\begin{aligned} p(\mathbf{X}_k|\mathcal{Z}^k) &\propto p(\mathbf{X}_k|\mathcal{Z}^{k-1}) \cdot \ell(\mathbf{X}_k|\mathcal{Z}_k) \\ &\propto \mathcal{W}(\mathbf{X}_k; n_{k|k-1}, \mathbf{C}_{k|k-1}) \cdot \mathcal{W}(\mathbf{X}_k; n_z, \mathbf{C}_z) \\ &\propto |\mathbf{X}|^{\frac{n_z+n_{k-1}-d-1-d-1}{2}} \cdot e^{-\frac{1}{2}tr((\mathbf{C}_{k|k-1}(\mathbf{C}_z+\mathbf{C}_{k|k-1})^{-1}\mathbf{C}_z)^{-1}\mathbf{X})}. \end{aligned} \quad (4.77)$$

¹The parameters of the measurement likelihood n_z and $\mathbf{C}_z$, always refer to the current measurement at time k .

From the multiplication result, it can be seen that the resulting posterior density is a Wishart density with parameters $n_{k|k}$ and $\mathbf{C}_{k|k}$, given by:

$$n_{k|k} = n_z + n_{k-1} - d - 1 \quad (4.78)$$

$$\mathbf{C}_{k|k} = \mathbf{C}_{k|k-1}(\mathbf{C}_z + \mathbf{C}_{k|k-1})^{-1}\mathbf{C}_z. \quad (4.79)$$

Since the Wishart distribution is normalized, solving the integral in the denominator in (4.73) is not required. Thus, It is sufficient to write the Wishart distributed posterior as:

$$\begin{aligned} p(\mathbf{X}_k | \mathcal{Z}^k) &= \mathcal{W}(\mathbf{X}_k; n_{k|k}, \mathbf{C}_{k|k}) \\ &= \frac{|\mathbf{X}|^{\frac{n_{k|k}-d-1}{2}}}{2^{\frac{n_{k|k}d}{2}} \Gamma_d(\frac{n_{k|k}}{2}) |\mathbf{C}_{k|k}|^{\frac{n_{k|k}}{2}}} e^{-\frac{1}{2} \text{tr}(\mathbf{C}_{k|k}^{-1} \mathbf{X})} \end{aligned} \quad (4.80)$$

The multiplication steps are shown in greater detail in Appendix B.1.

Prediction

In the prediction step, the parameters $n_{k|k-1}$ and $\mathbf{C}_{k|k-1}$ which define the predicted density

$$p(\mathbf{X}_k | \mathcal{Z}^{k-1}) = \mathcal{W}(\mathbf{X}_k; n_{k|k-1}, \mathbf{C}_{k|k-1}) \quad (4.81)$$

have to be calculated from $n_{k-1|k-1}$, $\mathbf{C}_{k-1|k-1}$ available after the previous filtering step. Following the same heuristic approach to predict the random matrix in [25], that is, the expectation of the predicted density shall be equal to the expectation of the previous filtering step—i.e., the target shape is unchanging over time, one can write:

$$\begin{aligned} \mathbb{E}[\mathbf{X}_k | \mathcal{Z}^{k-1}] &= \mathbb{E}[\mathbf{X}_{k-1} | \mathcal{Z}^{k-1}] \\ n_{k|k-1} \mathbf{C}_{k|k-1} &= n_{k-1|k-1} \mathbf{C}_{k-1|k-1} \\ \mathbf{C}_{k|k-1} &= \frac{n_{k-1|k-1} \mathbf{C}_{k-1|k-1}}{n_{k|k-1}}. \end{aligned} \quad (4.82)$$

Here the parameter n represents the degree of freedom or the sample size. Thus, it is directly related to the uncertainty of the corresponding expectation. Further, the uncertainty of the expectation shall be increasing at the prediction step in proportion to the update interval $\Delta t = t_k - t_{k-1}$. Another factor that needs to be considered,

is that the parameter n must satisfy $n \geq d$. After taking into account these factors and incorporating temporal decay constant τ as an additional modeling parameter. A plausible model to predict the parameter n , which incorporates a temporal decay constant τ as an additional modeling parameter could be:

$$n_{k|k-1} = d + e^{-\frac{\Delta t}{\tau}}(n_{k-1|k-1} - d) \quad (4.83)$$

The choice of the constant τ determines how agile the shape can be. In other words, decreasing τ increases the agility of the extent. An analogous model was introduced in [14] to predict the degree of freedom of the inverse Wishart density, and the models will be identical in case $d = 2$.

Unbiased Wishart Update

To check if our estimator is unbiased, we will verify if it meets the criteria in the definition of an unbiased estimator mentioned in [6].

Definition 4.3.2 : *An estimator $\hat{\mathbf{X}}$ is said to be unbiased if the expected value of the estimation error $\tilde{\mathbf{X}} \triangleq \mathbf{X} - \hat{\mathbf{X}}$ is a zero mean, that is,*

$$\mathbb{E}[\mathbf{X} - \hat{\mathbf{X}}] = 0. \quad (4.84)$$

Considering a point estimate $\hat{\mathbf{X}} = n\mathbf{C}$, the above unbiasedness requirement is satisfied if:

$$\mathbb{E}[\mathbf{X} - n\mathbf{C}] = 0. \quad (4.85)$$

By examining the requirement for unbiasedness in the derived posterior Wishart density parameters, one can demonstrate that the unbiased estimator of the Wishart density is given by:

$$\hat{\mathbf{X}} = n\mathbf{C} - n_z\mathbf{R}_d \quad (4.86)$$

where $\mathbf{R}_d$ is a $d \times d$ measurements error covariance matrix, the parameters n and $\mathbf{C}$ are calculated as:

$$n = n_z + n_0 \quad (4.87)$$

$$\mathbf{C} = \mathbf{C}_0(\mathbf{C}_z + \mathbf{C}_0)^{-1}\mathbf{C}_z \quad (4.88)$$

The full proof is provided in Appendix B.2.

Random Shape Contour Model

In this section, a model for a random shape contour applicable for LiDAR measurement is proposed. The contours of the shape are described by a set $\mathcal{R} = \{r^i\}_{i=1}^d$, for all d equidistantly discretisation in the interval $[0, \pi]$, which is convenient for representing abstract shapes. The goal is to estimate the length of each i radial segment r^i by the mean of the proposed Wishart filter. The key assumptions made in this model are the following:

- **Assumption.1.** The target extent is described by a *symmetric star convex* shape.
- **Assumption.2.** The kinematics's state and extent are decoupled, i.e., the target's centre and extent are tracked independently.
- **Assumption.3.** At each update frame, n_z measurements are associated with each radial segment.

In the Euclidean space, a set $\mathcal{S} \subseteq \mathbb{R}^n$ is called a star-convex set if the line segment which connects the centre $\mathbf{c}_k$ to any point $\mathbf{y}^i$ at the surface is fully contained in the set $\mathcal{S}$. Also, a shape is called symmetric if the two segments on the opposite side of $\mathbf{c}_k$ have equal length. In a Cartesian coordinate system, any point $\mathbf{y}^i \in \mathbb{R}^2$ lies on the contour of a star convex body corresponding to an angle θ^i can be obtained by

$$\begin{aligned} \mathbf{y}^i &= \mathbf{c}_k + \mathbf{H}^i r^i \\ &= \mathbf{c}_k + \begin{bmatrix} \cos(\theta^i) \\ \sin(\theta^i) \end{bmatrix} r^i \end{aligned} \quad (4.89)$$

where $\mathbf{c}_k$ is the target's centre and r^i is the radius' length which corresponds to the angle θ^i , Fig. 4.15.

Each measurement $\mathbf{z}^i$ stemming from one of the measurement sources along the contour is corrupted by a white Gaussian noise $\mathbf{u}_k^i$, characterized by a measurement error covariance matrix $\mathbf{R}$. Thus, the measurement equation can be written as

$$\begin{aligned} \mathbf{z}_k^i &= \mathbf{y}_k^i + \mathbf{u}_k^i \\ &= \mathbf{c}_k + \mathbf{H}^i r^i + \mathbf{u}_k^i \end{aligned} \quad (4.90)$$

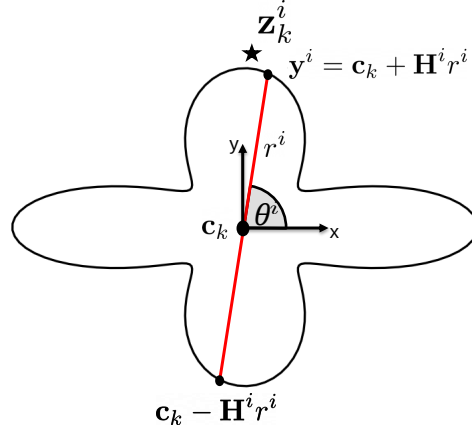


Fig. 4.15 Two measurement sources $\mathbf{H}^i r^i$ and $-\mathbf{H}^i r^i$ correspond to a single segment at angle θ^i of a symmetric star-convex shape. Measurements stemming from any of the two sources are associated with the segment in red.

and described by the Gaussian density

$$\mathbf{z}_k^i \sim \mathcal{N}(\mathbf{c}_k + \mathbf{H}^i r^i, \mathbf{R}) \quad (4.91)$$

The unknown state x^i which we aim to estimate is the segment's radius r^i . The value of r^i is not directly related to the measurement $\mathbf{z}^i$, but it has a specific trigonometric functional relation to the measurement at the corresponding angle (4.89). Because the trigonometric function has a constant value for each angle θ^i , the value of r^i can be obtained from either one of the measurement components z_x^i or z_y^i . Therefore, a sufficient measurement equation where the state r^i is completely observable could be:

$$z_x^i = c_x + r^i \cos(\theta^i) + u_x^i \quad (4.92)$$

or equivalently:

$$z_y^i = c_y + r^i \sin(\theta^i) + u_y^i \quad (4.93)$$

For brevity, The sub-index "x", "y" and the shift $\mathbf{c}$ are dropped in the rest of this section, and we will write the measurement equation of a measurement's source i as:

$$z^i = h^i r^i + u^i \quad (4.94)$$

Measurement Likelihood

As a result of symmetry (**Assumption.1**), a measurement z^i stemmed from any of the two ends of segment i corresponding to an angle θ^i , may be used to estimate the unknown r^i . Therefore, the measurement's distribution consists of two Gaussian densities, each of which corresponds to a different endpoint of segment i , and can be expressed by a Gaussian mixture of this form:

$$\begin{aligned} p(z^i|\theta^i) &= \frac{1}{2}p(z^i|h^i r^i) + \frac{1}{2}p(z^i|-h^i r^i) \\ &= \frac{1}{2}\mathcal{N}(z^i; h^i r^i, \sigma^2) + \frac{1}{2}\mathcal{N}(z^i; -h^i r^i, \sigma^2) \end{aligned} \quad (4.95)$$

According to **Assumption.2** the target centre is known, thus by shifting the measurements according to the tracked centre the measurements z^i will be centred around the coordinate's centre $(0,0)$. Accordingly we can establish that the probability density of the measurement is zero-mean. Based on the moments matching principle, the first moment of the mixture in (4.95) could be matched to a d-variate Gaussian distribution with unknown covariance matrix $\mathbf{C}$, that is,

$$\mathbf{z} = \begin{bmatrix} z^1 \\ \vdots \\ z^d \end{bmatrix} \sim \mathcal{N}(\mathbf{z}; 0, \mathbf{C}) \quad (4.96)$$

where each entry represents a single association to each radial segment $r^i \in \mathcal{R}$. Considering n_z realisations stemmed from the measurement's sources of each radial segment (**Assumption.3**), a similar data matrix to (4.67) can be constructed: $\mathbf{z} = [\mathbf{z}^1, \dots, \mathbf{z}^{n_z}]$. According to definition (1), the distribution of $d \times d$ random matrix $\mathbf{X} = \mathbf{z}\mathbf{z}^\top = \sum_{j=1}^{n_z} \mathbf{z}^j \mathbf{z}^{j\top}$ is a Wishart distribution. Consequently, the measurement likelihood of the random matrix $\mathbf{X}_k$ at time t_k given n_z association to each radial segment has a Wishart distribution:

$$\ell(\mathbf{X}_k|\mathbf{z}_k) = \mathcal{W}(\mathbf{X}_k; n_z, \mathbf{C}_z). \quad (4.97)$$

The measurement likelihood parameters are obtained as shown previously in (4.74). Considering a LiDAR type of sensor, only one side of the segments is observable in a single scan. This measurement characteristic compromises the correctness of the measurement likelihood in (4.95), which leads to undesirable results. In Appendix (B.3)

this behavior is discussed in more detail, also, it is shown that this effect is eliminated by assuming in-dependency between the segments.

On the grounds of this, adaptation to the measurement likelihood of the random matrix is required to ensure unbiasedness and to accommodate the structure of the shape in sufficient generality. To illustrate this effect, three models are tested and compared:

- General model:

The measurement of the random matrix is taken without any adaptation, thus the model assumes that the random matrix is calculated as:

$$\mathbf{X} = \mathbf{z}\mathbf{z}^\top \quad (4.98)$$

- Independent model:

In this model the radial segments are assumed to be independent, thus, the off-diagonal elements are set to zero:

$$\mathbf{X} = \text{diag}[\mathbf{z}\mathbf{z}^\top] \quad (4.99)$$

- Squared exponential:

The model encodes the assumption that two measurements of radial segments r and r' have higher correlation if their corresponding angles θ and θ' are closer to each other than segments far apart. Therefore, it is desirable to decrease the correlation (exponentially) across the angular distance. A common choice to represent this model is the squared exponential (SE) covariance function¹.

$$f(\theta, \theta') = e^{-\frac{|\theta - \theta'|^2}{2\alpha^2}} \quad (4.100)$$

where α is a constant determining the length of the covariance function. This can be incorporated into the measurement likelihood as:

$$\mathbf{X} = \mathbf{z}\mathbf{z}^\top \circ f(\theta, \theta') \quad (4.101)$$

where $\circ$ denotes the Hadamard product.

¹The squared exponential function is widely used in machine learning applications [48]. It was also used to model an analogous assumption in ETT contexts [47]

Evaluation Results

In this section, a numerical example is used to compare the performance of the biased with the unbiased estimators. Also, the proposed approach is tested and evaluated using simulations data, resembling the characteristics of a LiDAR sensor. Then the results are evaluated and compared with an alternative approach in the literature using the suggested metric.

Numerical Example

The effect of the unbiased estimator shall be illustrated by the following numerical example. Let a box object be centred at the origin $(0,0)$ with dimensions 4×6 . To estimate the dimensions of the box, a photogrammetric technique is utilised to measure the position of the edges' points/pixels in respect to the centre of the box, Fig.4.14.

Measurement likelihood At each frame k , the camera reports a number of noisy measurements $n_z = 10$ about the position of the edges' points/pixels. As a result of the symmetry, the reported measurement which stems from the boundary of the symmetric object is used to estimate the object radial segments r^1 and radius r^2 . Accordingly the measurement likelihood is expressed by the mixture density:

$$p(\mathbf{z}|\mathbf{r}) \sim \frac{1}{2}\mathcal{N}(\mathbf{z}; \mathbf{r}, \mathbf{R}) + \frac{1}{2}\mathcal{N}(\mathbf{z}; -\mathbf{r}, \mathbf{R}), \quad (4.102)$$

where $\mathbf{r} = [r^1 = 2, r^2 = 3]^\top$ and $\mathbf{R} = 0.25\mathbf{I}_{2 \times 2}$.

Consider the following measurement matrix at frame $k = 1$

$$\begin{aligned} \mathbf{z}_1 &= \begin{bmatrix} z_{k=1}^{1,1} & \cdots & z_{k=1}^{1,10} \\ z_{k=1}^{2,1} & \cdots & z_{k=1}^{2,10} \end{bmatrix} \\ &= \begin{bmatrix} -1.4 & -1.9 & -1.8 & 1.9 & 1.6 \\ -2.4 & 3.1 & -3.1 & -2.8 & 2.9 \\ -2.4 & -2.1 & 2.9 & -1.8 & -2.6 \\ -3.6 & -3.6 & -3.6 & 2.9 & 2.9 \end{bmatrix}. \end{aligned} \quad (4.103)$$

From definition (1), the parameters of the measurement likelihood

$$\ell(\mathbf{X}_1|\mathbf{z}_1) \sim \mathcal{W}(\mathbf{X}_1; n_z, \mathbf{C}_z), \quad (4.104)$$

are calculated as:

$$\begin{aligned} n_z &= 10. \\ \mathbf{C}_z &= \mathbf{z}_1 \mathbf{z}_1^\top / n_z \\ &= \begin{bmatrix} 4.3 & -0.5 \\ -0.5 & 9.9 \end{bmatrix}, \end{aligned} \quad (4.105)$$

the calculation of the likelihood follows the general model (4.98). In case of the independent model (4.99), the off-diagonal elements in $\mathbf{z}_1 \mathbf{z}_1^\top$ must be set to zero.

Prior The initial guess of the Wishart prior's parameters are set as

$$p(\mathbf{X}_0) = \mathcal{W}(\mathbf{X}_0; n_0, \mathbf{C}_0) \quad (4.106)$$

where $\mathbf{C}_0 = 15 \cdot \mathbf{I}_{2 \times 2}$ and $n_0 = 3$. Note that, for two dimensions the prior $n_0 = 2$ implies maximum uncertainty.

Posterior By the mean of the Wishart filter update shown in (B.6) and (B.7), the prior $p(\mathbf{X}_0)$ and the measurement likelihood $\ell(\mathbf{X}_1|\mathcal{Z}_1)$ are fused to obtain the posterior density at time $k = 1$

$$p(\mathbf{X}_1|\mathcal{Z}^1) = \mathcal{W}(\mathbf{X}_1; n_1, \mathbf{C}_1) \quad (4.107)$$

where

$$\begin{aligned} n_1 &= 13. \\ \mathbf{C}_1 &= \begin{bmatrix} 3.4 & -0.2 \\ -0.2 & 5.9 \end{bmatrix} \end{aligned}$$

By recursively update the posterior with new measurements $\mathbf{z}_k$, the obtained parameters of the posterior density after 6 update iterations are

$$p(\mathbf{X}_6|\mathcal{Z}^6) = \mathcal{W}(\mathbf{X}_1; n_6, \mathbf{C}_6) \quad (4.108)$$

where

$$n_6 = 63.$$

$$\mathbf{C}_6 = \begin{bmatrix} 0.6 & -0.06 \\ -0.06 & 1.2 \end{bmatrix}$$

To get a meaningful value of the Wishart density, the expected random matrix at frame $k = 6$ is calculated as:

$$\mathbb{E}[\mathbf{X}] = n_6 \mathbf{C}_6, \quad (4.109)$$

this matrix represent the expected value of the n_z sum of $\mathbf{r}$ squared, thus our estimate of the $\mathbf{r}$ is obtained as:

$$\begin{aligned} \mathbb{E}[\mathbf{r}] &= \text{diag} \left[\left(\frac{\mathbb{E}[\mathbf{X}]}{n_z} \right)^{-\frac{1}{2}} \right] \\ &= [2.0, 2.8]^\top \end{aligned} \quad (4.110)$$

where the term $()^{-\frac{1}{2}}$ is calculated using Cholesky decomposition. It can be seen in the results above (4.110), the estimates of r^1 and r^2 after only 6 iteration and under high noise scenario are very close to the really value.

Table 4.2 Root-mean-square error of different point estimators are listed in the table. The numbers are averaged over 1000 MC runs. The parameters n and $\mathbf{C}$ are calculated as in (B.6) and (B.7).

$\mathbb{E}[r]$	$RMSE(r^1 - \hat{r}^1)$	$RMSE(r^2 - \hat{r}^2)$
$n\mathbf{C}$	0.311	0.437
$n\mathbf{C} - n_z \mathbf{R}_d$	0.247	0.364

The selection of high noise measurement in this examples is made to allows a better demonstration of the biasedness. The performance of the unbiased point estimator is compared to the biased version in Table 4.2. The comparison metric used is the *Root Mean Square Error* (RMSE), which calculates the square root of the average squared error of N run as:

$$RMSE(\bar{x}, \mathbb{E}[x]) = \sqrt{\frac{1}{N}(\bar{x} - \mathbb{E}[x])^2}. \quad (4.111)$$

The number of Monte Carlo (MC) runs used in this simulation is $N = 1000$ MC. One can see that, the point estimator which subtracts the term $n_z \mathbf{R}_d$ has smaller RMSE comparing to the other.

Alternative approach

The proposed approach is compared with Gaussian Processes (GP) model for extended target tracking [47]. Similar to the proposed approach in this work, GP describes the target extent by a set of radial functions/segments represented by the vector $\mathbf{x}$;

$$\mathbf{x}^f = [f^1, f^2, \dots, f^d]^\top. \quad (4.112)$$

Where $\mathbf{x}^f$ is modeled as a GP, completely defined by a mean function $\mu(u)$ and a kernel function $k(\theta, \theta')$:

$$f(\theta) \sim \mathcal{GP}(\mu(u), k(\theta, \theta')) \quad (4.113)$$

The authors explicitly define the recursions as a state space model that enabled the use of a Kalman filter to update the posterior density of $f(\theta)$ using a time series of measurements. This results in the following recursion

$$p(\mathbf{f}|z_1 : N) \propto p(z_N|\mathbf{f}, z_{1:N-1})p(\mathbf{f}|z_{1:N-1}) \quad (4.114)$$

The underline measurement equation used in GP describes noisy measurements originating from the target contours of a *star-convex* shape. In order to model the periodicity in $f(\theta)$, a periodic kernel function $k(\theta, \theta')$ is used. This allows the incorporation of the symmetry assumption easily in the model by modifying the period of the kernel function to have a period of π rather than 2π . Against this background, using the GP approach as a benchmark is a suitable choice providing an objective comparison to our proposed approach.

Simulation

Different tests are performed to evaluate the proposed approach. The generated measurement points used in the tests simulate LiDAR sensor measurements originating from a measurement's sources along the boundary of an extended target. The first test aims to demonstrate the affect of the correlation in the Wishart measurement likelihood, which is discussed earlier in section 4.3.3. The target shape of choice in this test is a simple diamond shape. To capture the details for a shape of this type, it is sufficient to use only two radial segments. Thus the random matrix is represented by; $\mathbf{X}_{2 \times 2}$. With this choice of the target's shape and random matrix size, it becomes easier to demonstrate the effects of correlation, as only two radial segments are used. In Fig. 4.16, the results of the general and independent measurement models (4.98) and (4.99), respectively, are shown by the green and red shapes. The results demonstrate

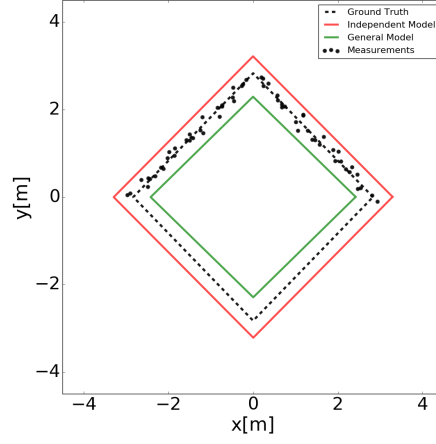
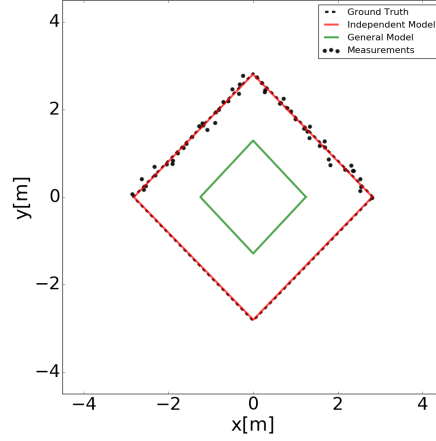
(a) $k = 2$ (b) $k = 10$

Fig. 4.16 Illustration of the estimated shapes by Wishart filter, using general and independent measurement models.

that the general measurement model underestimates the actual extent of the target. This undesirable result is directly related to the joint distribution of the measurement likelihood in the general model, which fails to capture the true distribution of the measurements when only one side of the segment is observable. Alternative models are suggested earlier to eliminate this effect, namely, the SE model in (4.101) and the independent model in (4.99). For this specific setup where only two far apart segments are used, the independent model and SE model are equivalent. For this reason, the SE model is discarded in this test, and only the independent model is compared to the general model. As expected, the independence assumption in the model yields better

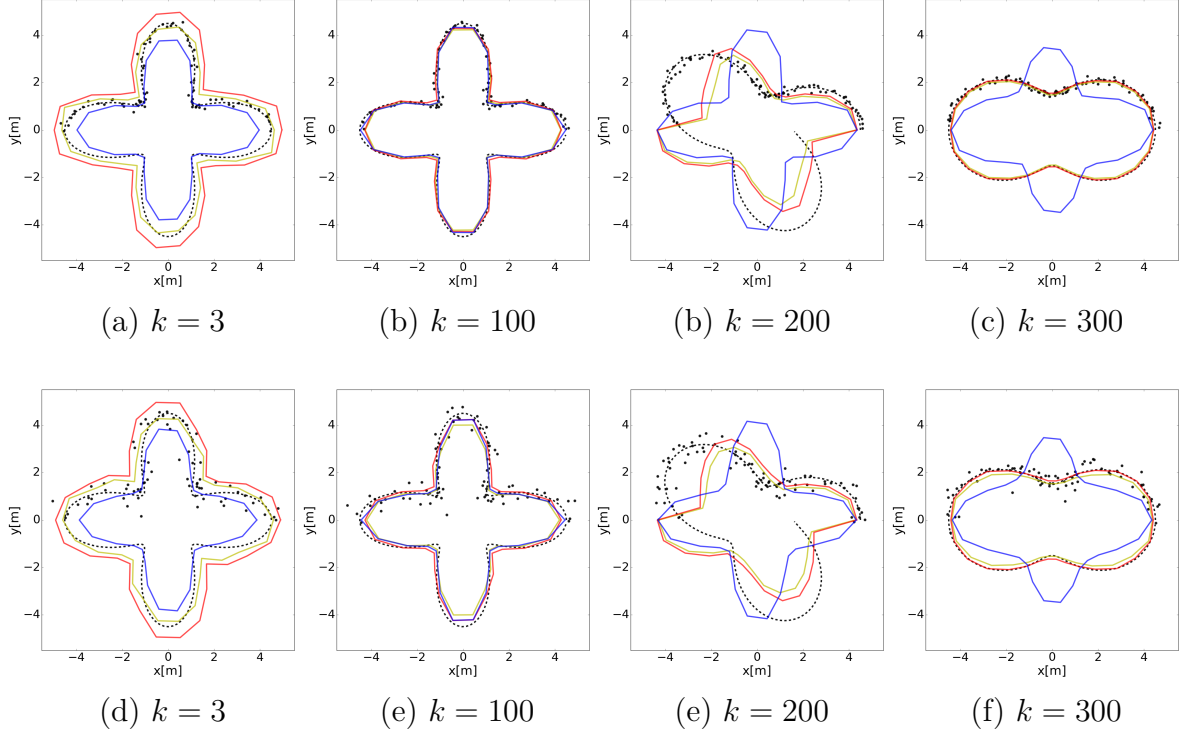


Fig. 4.17 In this simulation result, the extended target ground truth extent is represented by the black dashed shape and the dark blue point measurements resemble the LiDAR sensor reflection of one side of the target contours. The red and the yellow solid curves represent the estimated target extent by the proposed approach, corresponding to the independent and SE models. The GP approach estimated extent is represented by the blue solid curve.

results, where a very accurate estimation of the extent is obtained after few updates. This is shown by the red shape in Fig. 4.16.

The second test aims to compare the proposed Wishart filter with the chosen alternative approach (GP). In total, 400 frames of simulated measurement points are used in this test. At each frame, a random number of measurements are generated from a Poisson distribution, with an expected rate of occurrences $\lambda = 100$ and time interval between each scan as $\Delta t = 1ms$. In order to effectively evaluate the robustness of the proposed approach, a stationary target with a dynamic shape is used in the simulation. For the first half of the simulation frames, the algorithms are tested on a cross-section shape with smooth edges. In the second half of the frames, the target dynamically changes its shape to a less complex one. This allows us to test the robustness of the algorithm in handling new information regarding the target shape. This is encountered in real applications, due to a changes in sensor-to-target geometry and previously

occluded target. The accuracy of the algorithm is tested in two different noise scenarios, a low noise scenario and a high noise scenario, corresponding to measurement error covariance $\mathbf{R} = 0.09\mathbf{I}_{2 \times 2}$ and $\mathbf{R} = 0.01\mathbf{I}_{2 \times 2}$ respectively.

The results¹ of Wishart filter compared to the GP approach are illustrated in Fig. 4.17. In order to show the evolution of the estimated shapes, four snapshots are taken from different frames. The results in Fig. 4.17 demonstrate that the Wishart and GP models produce very accurate estimates of the shape in both low and high noise scenarios. However, when the target extent became dynamic (around frame $k = 200$), the Wishart filter appears to be more robust and capable of adapting to new shapes, while the GP model requires a relatively larger number of frames to adapt to the new shape. Furthermore, the agility of the estimated Wishart filter's shape directly related to the constant parameter τ of the Wishart prediction step, where the value of τ in (4.83) is chosen empirically and is set to $\tau = 0.01$ for both squared exponential and independent measurement models. This value is found to be a good trade-off between stability and agility of the estimated shape. On the other hand, the shape agility in the GP model is related to the process and the measurement noise parameters. Similarly, the values of these parameters are tuned empirically. Therefore and for the fidelity of the comparison, the parameters of both Wishart and GP models are chosen to be the best trade-off between accuracy and agility. The discretization level of the angle interval is a parameter choice representing the level of detail the estimated shape required to capture. In this test, the Wishart filter uses 15 radial segment to describe the shape, thus, the size of the random matrix $\mathbf{X}_{15 \times 15}$ with initial value $\mathbf{X}_0 = 15\mathbf{I}_{15 \times 15}$. To achieve the same level of details in the GP model, and despite of the symmetric assumption, the number of the radial segments must be double the number needed in the Wishart model. Thus the shape is represented by a state vector $\mathbf{x}_{30 \times 1}^f$ and a covariance matrix $\mathbf{P}_{30 \times 30}$ and initialized as $\mathbf{x}_0^f = 0.9\mathbf{I}_{30 \times 1}$ and $\mathbf{P}_0 = 0.1e^{\frac{-\sin^2(|\theta - \theta'|)}{2(\frac{\pi}{15})^2}}$.

Evaluation metric

The performance of the proposed approach is evaluated based on an *Intersection over Union* (IoU) metric. Which measures the accuracy of the estimated target's extent. In order to apply the IoU metric, the areas covered by the target ground truth extent A and the estimated target's extent $\hat{A}$ has to be calculated. Then the IoU is calculated as follows:

$$\text{IoU} = \frac{A \cap \hat{A}}{A \cup \hat{A}} \in [0, 1] \quad (4.115)$$

¹Video with estimated shapes.

Thus, IoU computes the ratio of the intersection over the union of the ground truth and the estimated extents. The ratio will be close to 1 in case the areas A and $\hat{A}$ are heavily overlapping, or close to 0 otherwise.

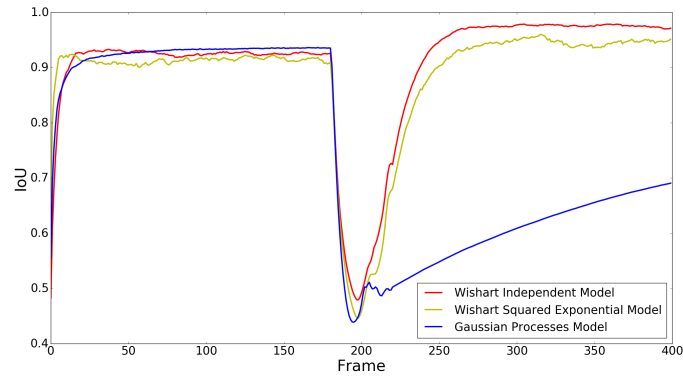
For the first half of the simulation the target maintains a cross-section shape. During this period, the IoU results illustrated in Fig. 4.18, show that the accuracy of the estimated shapes are almost similar. A slightly better performance is demonstrated by the GP model in the low noise scenario. The opposite is true for the high noise scenario, where the Wishart independent model performs slightly better than the other models. The robustness of the Wishart filter was clearly demonstrated in the second half of the simulation. In this, the target changed its shape to less a sophisticated shape and maintained this shape for the rest of the simulation. At the beginning of this period, the accuracy of the estimated shape dropped sharply for all the models. The Wishart filter models however, were able to adopt to the new shape, therefore, their corresponding IoU scores quickly recovered. Whereas, the GP model appears to require much longer time to capture the new shape. The behaviour of the GP model can be explained due to the convergence/shrinking of the state covariance over time. This behaviour doesn't present itself in the beginning because the initial covariance is relatively larger.

4.3.4 Conclusion

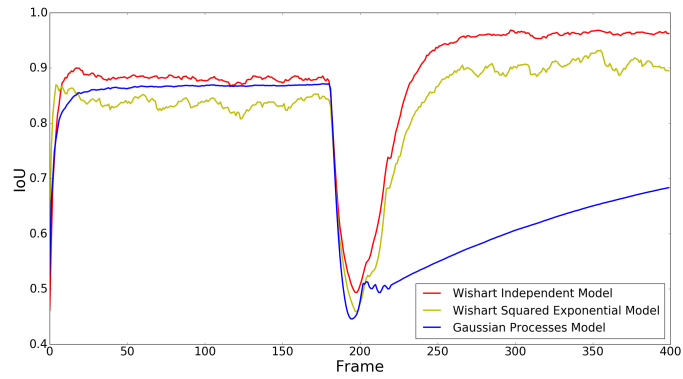
This section presented a state estimation problem where the state is known to take only positive real values. The application of choice to demonstrate this problem was the extent estimation problem in ETT, where the kinematics's state and the extent parameters are decoupled, i.e., treated independently. In this context, the section discussed the drawbacks of modeling the distribution of unknown quantities, i.e., the target extent, as Gaussian distributed random variables. Additionally, an alternative probability density limited to the positive real axis is suggested to model the target's extent, namely, a Wishart distribution.

A *symmetric star-convex* shape model is used to describe the target extent. Even though the symmetric assumption might sound restrictive, it is very well suited for many applications where the target to be tracked is symmetric, e.g., vehicle, motorcycle, aircraft, boat,...etc. Under the symmetric assumption in the shape, the possibility of modeling the measurements likelihood as a zero-mean Gaussian is discussed. This allows us to model the measurement likelihood as a Wishart density.

Considering a Wishart prior, we found that the Bayesian posterior density is also a Wishart distributed. Thus, the Wishart distribution is a conjugate prior to itself.



(a). Low noise scenario



(b). High noise scenario

Fig. 4.18 The figure illustrates the IoU results of the proposed Wishart filter for both independent and SE models shown in the red and yellow curves respectively. It is also compared to the GP approach shown by blue curve.

Consequently, an exact recursive Bayesian update formula for the Wishart distribution is obtained.

Although the target extent is assumed to be static, the prediction step plays an important role in the correctness of the uncertainty, also giving the shape the agility to adapt to new information. The simulation results show a better capability in the proposed approach to incorporate a new information compared to the chosen benchmark algorithm. In view of this, we may say that the proposed approach is more robust to difficult tracking scenarios where unseen parts of the target extent in previous sensor scans become visible due to factors such as occlusion and change in target-to-sensor geometry.

The general model incapacity to capture the joint distribution between the radial segments observed from only one side is discussed. Alternatively, it is proposed to incorporate our knowledge about the correlation in the measurement random matrix through the widely used squared exponential function. This model provides a convenient way to configure how strong the correlation between the segments is as a function of the distance. However, in practice the roughness of the target shape is unknown, which makes selecting a good value for the parameter α in the SE model a matter of guesswork. For instance, when the target to be tracked has a rough shape where the length of nearby segments is inconstant, the independent measurement model would be a better choice. Therefore, it can be said that the independent model provides more generality. The other design parameter which needs to be chosen is the number of segments d to describe the shape. To select the right level of details through setting up the value of d , one needs to consider the expected number of measurement points to be associated with each segment. That is because the Wishart distribution exists in case $n \geq d$. To state this differently, the fewer the measurement points reflected from the target, the less segments to describe the shape can be used.

4.3.5 Summary

The work presented in this dissertation primarily focused on extended and coordinated target tracking. Since group target tracking problems can fall under both extended and coordinated target tracking, and given that the work related to coordinated target tracking only considered the aspect of identifying coordination, group target tracking was addressed in the latter context. One can utilise approaches for the simultaneous estimation of shape and kinematics (extended target) and combine them with automatic detection methods for group targets. In this work, we assumed that the shape of the group is independent; however, it is worth investigating in the future whether it

is possible to infer more information about group behaviour based on the estimated shape.

At the time the author is writing this summery at the end of his PhD work, the author can argu that the avalible solutions presented in the avalaible solution in the litrature concering extended target tracking problem has reach a very good stage, many rubost, fast and accuraty algorithms can be deployed for prodcution however, there is still work to do. In the other hand the coordinated target problem is relatively unexplored area and the development of smart mulit-agent system will push the need for a system which can detect coordination and colleactive behivore ¹

The the presented work in this dissertation focused manly on extended and coordinated target. As group targets problem can fall within the both extended and coordinate target tracking and as the work related to coordinated target track is only considere the aspect of identifying coordination the group target tracking was addressed in the latter. On can utalize the approachs for simultiinous estimation of shape and kinematics (extended target) and combine it with automatic detection of group target approach sections collectively, our work assumed that the shape of the group is indpendent however, it is worth reseraching in the future, if it is possible to infare more information of the group behavior based on the estamted shape.

present a series of methods for improving target tracking by estimating both kinematic and shape variables simultaneously. The proposed approaches employ mathematical models and alternative probability distributions, such as Gaussian mixtures in the Extended Kalman Filter (EKF) and Wishart distributions for extent estimation, which are particularly suited for tracking symmetric objects like vehicles, boats, and aircraft. Despite certain limitations, such as difficulty in accurately modeling targets when only one side is visible, these methods show robust performance, especially in scenarios where parts of the target's extent become visible due to occlusions or changes in target-to-sensor geometry.

Furthermore, the research addresses challenges in modeling target shapes, suggesting methods to handle correlations between measurements and uncertainties using functions like the squared exponential. While these models provide flexibility, selecting appropriate parameters remains a challenge due to unknown target characteristics. The work also examines the problem of coordinated target tracking, enhancing the separation of closely spaced targets and identifying collaborative behaviors among multiple units using a Bayesian framework.

¹If the computer scientes solve this prooblem, the beiologist will be thankfull. In fact as result of publishing a our paper on coordinated target detection []

The findings demonstrate that by using adaptive state dynamics (ASDs) and applying scoring functions for group detection, the proposed methods offer improved accuracy and resilience compared to traditional tracking algorithms. However, practical considerations, such as balancing computational cost and performance, are critical, particularly in optimizing the tracking window size and refining models to handle diverse scenarios. Overall, these approaches offer significant advancements in tracking accuracy and robustness in complex environments.

Chapter 5

Coordinated Targets

In this chapter, we will discuss the additional target characteristics addressed in this dissertation, specifically group and coordinated targets. We draw a distinction between group target and coordinated target in section 2.2. Group target are a well-studied area in tracking literature [19, 34, 30], but one could argue that closely spaced targets in a group also exhibit a type of coordination. Therefore, we will discuss both characteristics in this chapter. Additionally, we will present our methods to enhance situational awareness in surveillance systems when dealing with group and coordinated targets.

5.1 Introduction

Advances in automation and artificial intelligence have enabled the creation of large distributed, interconnected, and cooperative agents, in particular, systems with groups of *Unmanned Air Systems* (UAS) or *Unmanned Surface Vessels* (USV) which can be created based on low-cost commercial off-the-shelf hardware and open-source software. Since navigation in the air or on the water is almost trivial, the barrier to employ such a collaborative attack is quite low, which yields a new threat in asynchronous warfare for instance. By means of more sophisticated algorithms for navigation in GPS-free environments [40], light-weight protection measures such as GPS jamming may also be circumvented by a potential attacker. This has led to disruptive challenges for defence and security applications such as *Anti-Access/Aerial Denial* (A2/AD) since such systems are traditionally focused on protecting against singular targets. As a consequence, potential effectors are easily saturated by a sufficiently large amount of group members. This saturation opens up a security hole as members of the group may enter the area while the effectors are busy with other incoming. Members of

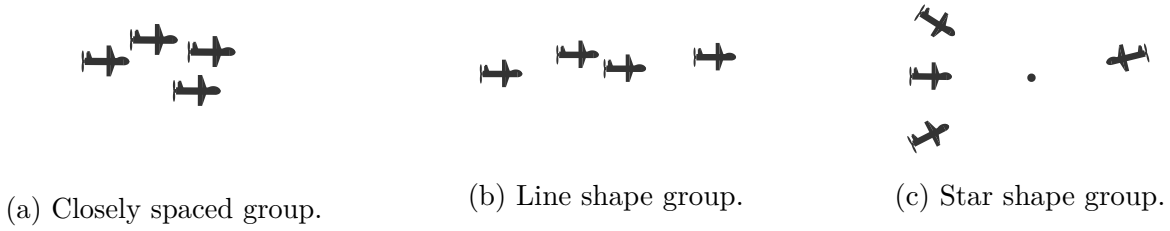


Fig. 5.1 Illustration of different coordination structures.

a coordinated group could also be used as decoys to distract the system from the actual targets and create confusion. Moreover, each member could be equipped with environmental sensors and artificial intelligent algorithms capable of identifying targets and defence systems, the information observed by a single member is shared with the rest of the group, which gives the group/swarm an advantage to coordinate more efficiently to avoid defences and accomplish their shared objective. Current surveillance systems suffer from scalability problems and do not appear sufficient, a U.S. Army study [49] showed that swarming reduces the effectiveness of current defence systems to at least 50%.

In this work three types of coordinated groups are distinguished:

- **Closely spaced:** the members of the group share common kinematics and members are closely spaced, where each member has at least one neighbour located within a certain radius.
- **Line shape:** the group members form a line shape, where members reach their target consecutively from the same angle at different times.
- **Star shape:** the group members reach their target from different angles and could arrive at the same or at different time.

The *closely spaced* target type can be handled similarly to the methods used in group-target tracking problems. A common approach is to detect groups using spatial clustering algorithms. In this chapter, we propose an improvement based on utilising accumulated state densities (ASDs) within the density-based spatial clustering of applications with noise (DBSCAN) algorithm. Additionally, we will address other types of coordinated targets, specifically *line shape* and *star shape*, where the assumptions that group members are closely spaced and share common kinematics are not required.

5.2 Background on Accumulated State Densities

In most tracking algorithms, the characteristics of conditional probability densities $p(\mathbf{x}_l | \mathcal{Z}^k)$ of (joint) object states $\mathbf{x}_l$ are calculated, which describe the available knowledge of the object properties at a certain instant of time t_p , given a time series $\mathcal{Z}^k$ of imperfect sensor data accumulated up to time t_k . In certain applications, however, the kinematic object states $\mathbf{x}_k, \dots, \mathbf{x}_n$, $n \leq k$, accumulated over a certain time window from a past instant of time t_n up to the present time t_k is of interest. The statistical properties of the accumulated state vectors are completely described by the joint probability density function (pdf) of them, $p(\mathbf{x}_k, \dots, \mathbf{x}_n | \mathcal{Z}^k)$ which is conditioned by the time series $\mathcal{Z}^k$. These densities may be called accumulated state densities (ASDs) [27]. By marginalising them, the standard filtering and retrodiction densities directly result; in other words, ASDs provide a unified description of filtering and retrodiction as shown schematically in Figure 1. In addition, ASDs fully describe the correlations between the state estimates at different instants of time [17].

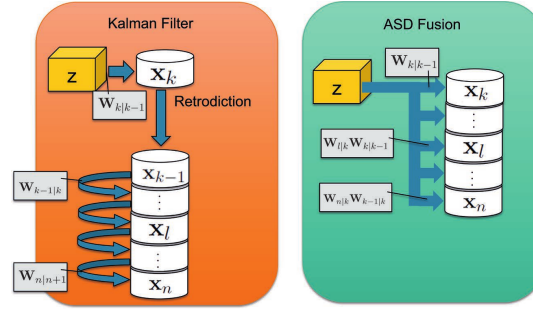


Fig. 5.2 Schematic comparison of ASD and Kalman filter fusion. Image courtesy of Govaers and Koch [17]

Originally, the Accumulated State Density (ASD) had been proposed to provide an exact solution to the out-of-sequence measurement problem, i.e. when the sensor data does not arrive in the temporal order in which they have been produced [27]. Another application of ASDs was demonstrated in [29] where it was shown that an exact solution for track-to-track fusion can also be achieved as a convex combination of local ASDs generated at each node in a distributed sensor system. The ASDs play a crucial role in the sets of trajectories (PHD for ASDs), where the authors in [15] presented an online estimator of a set of trajectories/tracks in a random finite set framework, that provides a minimal representation of the multiple target tracking system at all time steps.

5.2.1 Notion of ASD

All information on the object states accumulated over a time window $t_k, t_{k-1}, \dots, t_n$ of length $k - n + 1$,

$$\mathbf{x}_{k:n} = (\mathbf{x}_k, \dots, \mathbf{x}_n). \quad (5.1)$$

that can be extracted from the time series of accumulated sensor data $\mathcal{Z}^k$ up to and including time t_k is contained in a joint density function $p(\mathbf{x}_{k:n}|\mathcal{Z}^k)$, which may be called Accumulated State Density (ASD). Via marginalising over $\mathbf{x}_k, \dots, \mathbf{x}_{l+1}, \mathbf{x}_{l-1}, \dots, \mathbf{x}_n$,

$$p(\mathbf{x}_l|\mathcal{Z}^k) = \int d\mathbf{x}_k, \dots, d\mathbf{x}_{l+1}, d\mathbf{x}_{l-1}, \dots, d\mathbf{x}_n p(\mathbf{x}_k, \dots, \mathbf{x}_n|\mathcal{Z}^k), \quad (5.2)$$

the filtering density $p(\mathbf{x}_k|\mathcal{Z}^k)$ for $l = k$ and the retrodiction densities $p(\mathbf{x}_l|\mathcal{Z}^k)$ for $l < k$ result from the ASD. ASDs thus in a way unify the notions of filtering and retrodiction. In addition, ASDs also contain all mutual correlations between the individual object states at different instants of time. Bayes' Theorem provides a recursion formula for updating accumulated state densities:

$$p(\mathbf{x}_{k:n}|\mathcal{Z}^k) = \frac{p(\mathbf{Z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1:n}|\mathbf{Z}^{k-1})}{\int d\mathbf{x}_k p(\mathbf{Z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1:n}|\mathcal{Z}^{k-1})} \quad (5.3)$$

The sensor data $\mathbf{Z}_k$ explicitly appear in this representation. A little formalistically speaking, 'sensor data processing' means nothing else than to achieve by certain reformulations that the sensor data are no longer be explicitly present.

5.2.2 ASDs Closed-form Representation

Under conditions, where Kalman filtering is applicable (perfect data sensor-data-to-track association, linear Gaussian sensor and evolution models), a closed-form representation of $p(\mathbf{x}_{k:n}|\mathcal{Z}^k)$ can be derived. In this case, let the likelihood function be given by:

$$p(\mathbf{Z}_k|\mathbf{x}_k) = \mathcal{N}(\mathbf{z}_k; \mathbf{H}_k \mathbf{x}_k, \mathbf{R}_k), \quad (5.4)$$

where $\mathbf{Z}_k = \mathbf{z}_k$ denotes the vector of sensor measurements at time t_k , $\mathbf{x}_k = \mathbf{x}_k$ the kinematic state vector of the object, $\mathbf{H}_k$ the measurement matrix, and $\mathbf{R}_k$ the measurement error covariance matrix, while the Markovian evolution model of the target

is represented by:

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \mathbf{F}_{k|k-1} \mathbf{x}_{k-1}, \mathbf{D}_{k|k-1}), \quad (5.5)$$

with an evolution matrix $\mathbf{F}_{k|k-1}$ and a corresponding evolution covariance matrix $\mathbf{D}_{k|k-1}$. A repeated use of a well-known product formula for Gaussians [28] directly yields a product representation of the augmented state density:

$$p(\mathbf{x}_{k:n} | \mathcal{Z}^k) = \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k}, \mathbf{P}_{k|k}) \times \prod_{l=n}^{k-1} \mathcal{N}(\mathbf{x}_l; \mathbf{h}_{l|l+1}(\mathbf{x}_{l+1}), \mathbf{R}_{l|l+1}), \quad (5.6)$$

where the auxiliary quantities $\mathbf{h}_{l|l+1}, \mathbf{R}_{l|l+1}, l \leq k$, are defined by:

$$\mathbf{h}_{l|l+1}(\mathbf{x}_{l+1}) = \mathbf{x}_{l|l} + \mathbf{W}_{l|l+1}(\mathbf{x}_{l+1} - \mathbf{x}_{l+1|l}) \quad (5.7)$$

$$\mathbf{R}_{l|l+1} = \mathbf{P}_{l|l} - \mathbf{W}_{l|l+1} \mathbf{P}_{l|l+1} \mathbf{W}_{l|l+1}^\top \quad (5.8)$$

and a “retrodiction gain” matrix

$$\mathbf{W}_{l|l+1} = \mathbf{P}_{l|l} \mathbf{F}_{l+1|l}^\top \mathbf{P}_{l+1|l}^{-1}. \quad (5.9)$$

Note that $\mathcal{N}(\mathbf{x}_l; \mathbf{h}_{l|l+1}(\mathbf{x}_{l+1}), \mathbf{R}_{l|l+1})$ can be interpreted in analogy to a Gaussian likelihood function with a linear measurement function $\mathbf{h}_{l|l+1}(\mathbf{x}_{l+1})$. $\mathbf{h}_{l|l+1}, \mathbf{R}_{l|l+1}$ are defined by the parameters of $p(\mathbf{x}_l | \mathcal{Z}^l) = \mathcal{N}(\mathbf{x}_l; \mathbf{x}_{l|l}, \mathbf{P}_{l|l})$,

$$\mathbf{x}_{l|l} = \begin{cases} \mathbf{x}_{l|l-1} + \mathbf{W}_{l|l-1}(\mathbf{z}_l - \mathbf{H}_l \mathbf{x}_{l|l-1}) \\ \mathbf{P}_{l|l}(\mathbf{P}_{l|l-1}^{-1} \mathbf{x}_{l|l-1} + \mathbf{H}_l^\top \mathbf{R}_l^{-1} \mathbf{z}_l) \end{cases} \quad (5.10)$$

$$\mathbf{P}_{l|l} = \begin{cases} \mathbf{P}_{l|l-1} - \mathbf{W}_{l|l-1} \mathbf{S}_{l|l-1} \mathbf{W}_{l|l-1}^\top \\ (\mathbf{P}_{l|l-1}^{-1} + \mathbf{H}_l^\top \mathbf{R}_l^{-1} \mathbf{H}_l)^{-1} \end{cases} \quad (5.11)$$

Note that there exist two equivalent formulations of the Kalman update formulae according to the two versions of the product formula [28]. The innovation covariance matrix $\mathbf{S}_{l|l-1}$ and the Kalman Gain matrix are give by:

$$\mathbf{S}_{l|l-1} = \mathbf{H}_l \mathbf{P}_{l|l-1} \mathbf{H}_l^\top + \mathbf{R}_l. \quad (5.12)$$

$$\mathbf{W}_{l|l-1} = \mathbf{P}_{l|l-1} \mathbf{H}_{l|l-1}^\top \mathbf{S}_{l|l-1}^{-1}. \quad (5.13)$$

Also the parameters of the prediction density $p(\mathbf{x}_{l+1}|\mathcal{Z}^l) = \mathcal{N}(\mathbf{x}_{l+1}; \mathbf{x}_{l+1|l}, \mathbf{P}_{l+1|l})$,

$$\mathbf{x}_{l|l-1} = \mathbf{F}_{l|l-1} \mathbf{x}_{l-1|l-1} \quad (5.14)$$

$$\mathbf{P}_{l|l-1} = \mathbf{F}_{l|l-1} \mathbf{P}_{l-1|l-1} \mathbf{F}_{l|l-1}^\top + \mathbf{D}_{l|l-1}, \quad (5.15)$$

enter into the product representation in (5.6). With $\mathbf{x}_{l|k}$, $\mathbf{P}_{l|k}$, $\mathbf{W}_{l|l+1}$ known from the Rauch-Tung-Striebel recursion [36],

$$\mathbf{x}_{l|k} = \mathbf{x}_{l|l} + \mathbf{W}_{l|l+1}(\mathbf{x}_{l+1|k} - \mathbf{x}_{l+1|l}) \quad (5.16)$$

$$\mathbf{P}_{l|k} = \mathbf{P}_{l|l} + \mathbf{W}_{l|l+1}(\mathbf{P}_{l+1|k} - \mathbf{P}_{l+1|l}) \mathbf{W}_{l|l+1}^\top, \quad (5.17)$$

we can rewrite $p(\mathbf{x}_{k:n}|\mathcal{Z}^k)$ by the following product:

$$p(\mathbf{x}_{k:n}|\mathcal{Z}^k) = \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k|k}, \mathbf{P}_{k|k}) \prod_{l=n}^{k-1} \quad (5.18)$$

$$\mathcal{N}(x_l - \mathbf{W}_{l|l+1} x_{l+1}; \mathbf{x}_{l|k} - \mathbf{W}_{l|l+1} \mathbf{x}_{l+1|k}, \mathbf{Q}_{l|k})$$

where we used the abbreviation:

$$\mathbf{Q}_{l|k} = \mathbf{P}_{l|k} - \mathbf{W}_{l|l+1} \mathbf{P}_{l+1|k} \mathbf{W}_{l|l+1}^\top. \quad (5.19)$$

Due to elementary matrix algebra manipulations it can be shown that this product can be represented by a single Gaussian,

$$p(\mathbf{x}_{k:n}|\mathcal{Z}^k) = \mathcal{N}(\mathbf{x}_{k:n}; \mathbf{x}_{k:n}^k, \mathbf{P}_{k:n}^k) \quad (5.20)$$

with a joint expectation vector $\mathbf{x}_{k:n}^k$ defined by:

$$\mathbf{x}_{k:n}^k = (\mathbf{x}_{k|k}^\top, \mathbf{x}_{k-1|k}^\top, \dots, \mathbf{x}_{n|k}^\top)^\top, \quad (5.21)$$

while the corresponding joint covariance matrix $\mathbf{P}_{k:n}^k$ can be written as an inverse of a tridiagonal block matrix (see 5.26 on the subsequent page).

This can be seen by considering projectors Π_l defined by:

$$\begin{aligned}\Pi_l \mathbf{x}_{k:n} &= \begin{cases} (\mathbf{1}, \mathbf{0}, \dots, \mathbf{0}) \mathbf{x}_{k:n}, & l = k \\ (\mathbf{0}, \dots, -\mathbf{W}_{l|l+1}, \mathbf{1}, \dots, \mathbf{0}) \mathbf{x}_{k:n}, & n \leq l \leq k. \end{cases} \\ &= \begin{cases} \mathbf{x}_k, & l = k \\ \mathbf{x}_l - \mathbf{W}_{l|l+1} \mathbf{x}_{l+1}, & n \leq l \leq k. \end{cases}\end{aligned}\quad (5.22)$$

Using Π_l and $\mathbf{Q}_{l|k}$, $l = 1, \dots, k$, the ASD can be rewritten:

$$\begin{aligned}p(\mathbf{x}_{k:n}^k | \mathcal{Z}^k) &= \prod_{l=n}^k \mathcal{N}(\Pi_l \mathbf{x}_{k:n}; \Pi_l \mathbf{x}_{k:n}^k, \mathbf{Q}_{l|k}) \\ &= \prod_{l=n}^k \mathcal{N}(\mathbf{x}_{k:n}; \mathbf{x}_{k:n}^k, (\Pi_l^\top \mathbf{Q}_{l|k}^{-1} \Pi_l)^{-1}) \\ &= \mathcal{N}(\mathbf{x}_{k:n}; \mathbf{x}_{k:n}^k, \mathbf{P}_{k:n}^k)\end{aligned}\quad (5.23)$$

with a covariance matrix $\mathbf{P}_{k:n}^k$, which is given by an harmonic mean according to the product formula for Gaussians:

$$\mathbf{P}_{k:n} = \left(\sum_{l=n}^k \Pi_l^\top \mathbf{Q}_{l|k}^{-1} \Pi_l \right)^{-1}. \quad (5.24)$$

The summation of the matrices $\Pi_l^\top \mathbf{Q}_{l|k}^{-1} \Pi_l$ directly yields the inverse ASD covariance matrix as a tridiagonal block matrix displayed on the top of page 4. Here, the auxiliary quantities $\mathbf{T}_{l|k}$, $n \leq l \leq k$ are defined by:

$$\mathbf{T}_{l|k} = \begin{cases} \mathbf{Q}_{n|k}^{-1} & \text{for } l = n \\ \mathbf{P}_{k|k}^{-1} + \mathbf{W}_{l-1|l}^\top \mathbf{Q}_{l-1|k}^{-1} \mathbf{W}_{l-1|l} & \text{for } l = k \\ \mathbf{Q}_{l|k}^{-1} + \mathbf{W}_{l-1|l}^\top \mathbf{Q}_{l-1|k}^{-1} \mathbf{W}_{l-1|l} & \text{else} \end{cases} \quad (5.25)$$

The tridiagonal structure is a consequence of the Markov property of the underlying evolution model. This representation of the inverse of $\mathbf{P}_{k:n}^k$ is useful in calculations. By a repeated use of the matrix inversion lemma and an induction argument, the inverse of this tridiagonal block matrix can be calculated. The resulting block matrix is

$$\mathbf{P}_{k:n}^k = \begin{bmatrix} \mathbf{T}_{k|k} & -\mathbf{W}_{k-1|k}^\top \mathbf{Q}_{k-1|k}^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ -\mathbf{Q}_{k-1|k}^{-1} \mathbf{W}_{k-1|k} & \mathbf{T}_{k-1|k} & -\mathbf{W}_{k-2|k}^\top \mathbf{Q}_{k-2|k}^{-1} & \ddots & \vdots \\ \mathbf{0} & -\mathbf{Q}_{k-2|k}^{-1} \mathbf{W}_{k-2|k} & \ddots & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \mathbf{T}_{n+1|k} & -\mathbf{W}_{n|k}^\top \mathbf{Q}_{n|k} \\ \mathbf{0} & \cdots & \mathbf{0} & -\mathbf{Q}_{n|k} \mathbf{W}_{n|k} & \mathbf{T}_{n|k} \end{bmatrix}^{-1} \quad (5.26)$$

$$= \begin{bmatrix} \mathbf{P}_{k|k} & \mathbf{P}_{k|k} \mathbf{W}_{k-1|k}^\top & \mathbf{P}_{k|k} \mathbf{W}_{k-2|k}^\top & \cdots & \mathbf{P}_{k|k} \mathbf{W}_{n|k}^\top \\ \mathbf{W}_{k-1|k} \mathbf{P}_{k|k} & \mathbf{P}_{k-1|k} & \mathbf{P}_{k-1|k} \mathbf{W}_{k-2|k-1}^\top & \ddots & \mathbf{P}_{k-1|k} \mathbf{W}_{n|k-1}^\top \\ \mathbf{W}_{k-2|k} \mathbf{P}_{k|k} & \mathbf{W}_{k-2|k-1} \mathbf{P}_{k-1|k} & \mathbf{P}_{k-2|k} & * & \vdots \\ \vdots & * & * & * & \mathbf{P}_{n+1|k} \mathbf{W}_{n|n+1}^\top \\ \mathbf{W}_{n|k} \mathbf{P}_{k|k} & \mathbf{W}_{n|k-1} \mathbf{P}_{k-1|k} & \cdots & \mathbf{W}_{n|n+1} \mathbf{P}_{n+1|k} & \mathbf{P}_{n|k} \end{bmatrix} \quad (5.27)$$

displayed in (5.27), where the following abbreviations were used:

$$\mathbf{W}_{l|k} = \prod_{\lambda=l}^{k-1} \mathbf{W}_{\lambda|\lambda+1} = \prod_{\lambda=l}^{k-1} \mathbf{P}_{\lambda|\lambda} \mathbf{F}_{\lambda+1|\lambda}^\top \mathbf{P}_{\lambda+1|\lambda}^{-1} \quad (5.28)$$

The densities $\{\mathcal{N}(\mathbf{x}_l; \mathbf{x}_{l|k}, \mathbf{P}_{l|k})\}_{l=n}^k$ are directly obtained via marginalising, since the covariance matrices $\mathbf{P}_{l|k}$, $n \leq l \leq k$, appear on the diagonal of this block matrix. Note that the ASD is completely defined by the results of prediction, filtering, and retrodiction obtained for the time window $t_k, \dots, t_n$, i.e. it is a by-product for Kalman filtering and Rauch-Tung-Striebel smoothing.

5.3 Accumulated State Densities Filter for Better Separability of Group-Targets

In some defence applications, it is required to identify targets separated by a certain distance as group-targets. This allows the system to use a suitable tracking and mitigation strategy for a group different from what is used for a point-target. A natural choice to identify a group of this type is the DBSCAN algorithm. The DBSCAN algorithm uses the available track information to identify the groups/clusters. Information on these tracks are, in the vast majority of tracking systems, based on the

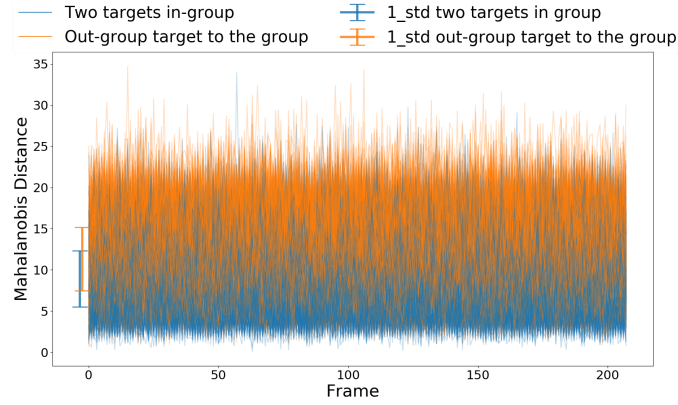
Kalman filter estimate. In this work, we present a scenario where an out-group target is indistinguishable from a group target when using a Kalman filter. We then demonstrate that the separability can be significantly enhanced using ASDs, which are estimates of the joint probability density of the kinematic target states accumulated over a certain time window up to the present time, based on the time series of all sensor data. In this approach, ASDs improve target separability by utilising the Mahalanobis distance between the available tracks in situations where the tracks generated by Kalman filters are inseparable. Using track information offers certain advantages over relying solely on measurement data. Furthermore, in some applications, such as distributed sensor networks, we are often limited to using track information since the measurements are not transmitted across the network.

5.3.1 Spatial Clustering Method

To identify a group-target with an arbitrary shape separated by a bounded distance, DBSCAN [11] is a natural choice, as one of its two parameters ϵ allows to specify the required distance bound between two points (or targets in our case) to be considered neighbours. While the other parameter *MinPts* specifies the minimum number of neighbour-points required to consider the point as dense and therefore identify a cluster. Moreover, the DBSCAN algorithm doesn't require to specify the number of clusters as an input parameter, instead, DBSCAN can discover any number of clusters. Additionally, the algorithm can identify clusters of varying size and shape and also it's not required that every point is assigned to a cluster but instead considers the far sparse points as outliers (or a point-targets in our case). The DBSCAN starts by transforming the space according to the density of the points, while points are separated as dense points and sparse points. Then, a linkage clustering is done based on the parameters ϵ and *MinPts* results in a dendrogram, then any singleton is considered to be noise.

5.4 Target State Separability

The quality of the DBSCAN depends on the distance metric used in the algorithm. In target tracking applications where the covariance information of the estimated target's state is available, the most common distance metric used is the Mahalanobis distance, as it provides an elegant mean to incorporate the confidence about the target state estimate. Based on the Mahalanobis distance, the distance metric between track $\mathbf{x}^a$



(a) Kalman Filter

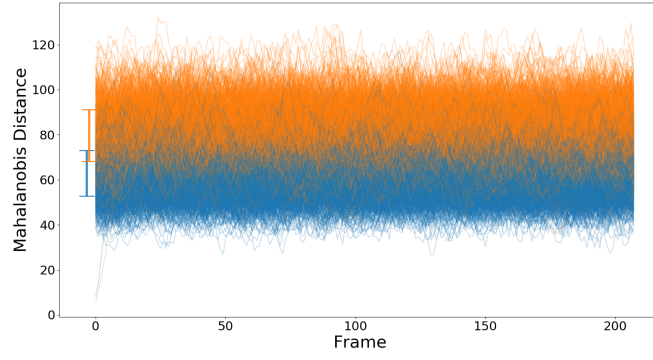
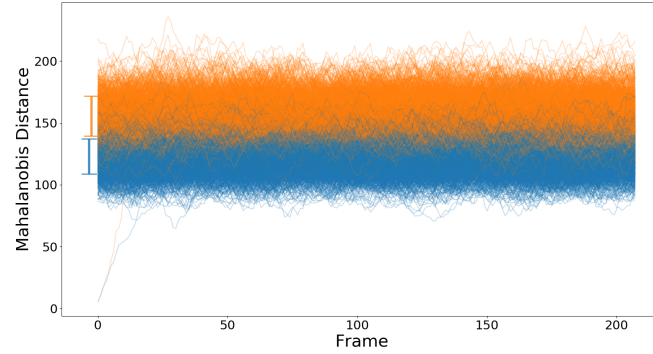
(b) ASDF_{n=10}(c) ASDF_{n=20}

Fig. 5.3 This figure shows the Mahalanobis distance obtained by 1000 Monte Carlo run of targets tracked using different state density sizes, where the total number of tracking frames is 200. The blue colour corresponds to targets separated by $50m$, while the orange colour corresponds to targets separated by $65m$. The error bar correspond to one standard deviation of the Mahalanobis distance.

and track $\mathbf{x}^b$ is provided by:

$$d(\mathbf{x}_k^a, \mathbf{x}_k^b) = (\mathbf{x}_k^a - \mathbf{x}_k^b)^\top (\mathbf{P}_k^a + \mathbf{P}_k^b)^{-1} (\mathbf{x}_k^a - \mathbf{x}_k^b), \quad (5.29)$$

based on this distance metric, the parameter ϵ must be tuned to meet the requirements of a particular application, such that $d(\mathbf{x}_k^a, \mathbf{x}_k^b) \leq \epsilon$. However, finding an appropriate value for ϵ is quite difficult. In fact, due to the stochastic nature of the estimates, the tracks can become easily inseparable when they get close to each other (relative to the value of ϵ). To illustrate this, let's consider three targets moving in linear trajectory, the distance between the members of the group is $50m$, while the distance between other non-member targets to the nearest group member is $65m$. The targets' Cartesian positions are measured using a noisy sensor — each measurement stemming from one of the targets is corrupted by white Gaussian noise characterised by a covariance matrix $\mathbf{R} = \text{diag}[12^2, 12^2]$. Under perfect data sensor-data-to-track association, the measurements are feed to a Kalman filter with constant velocity evolution model to track the target position. Fig. 5.3 a, shows the squared Mahalanobis distance of 1000 Monte Carlo run of the complete course of the simulation scenario, between two members group-target and between out-group point-target to the nearest member of the group-target in blue and orange color, respectively. One can easily see that, the group-target members are inseparable from the other out-group point-target in this scenario. However, the separability clearly improves by using ASDs in proportion to the chosen time window size, Fig 5.3 b and c.

5.4.1 Simulation and Results

Considering a defence system able to mitigate group-target within a bounded distance d between each member in the group and it's nearest neighbour. The goal is to identify a group-targets under distance bound constraints. Therefore, a scenario simulating such a constraint is implemented to evaluate the effectiveness of the ASDF on the targets' separability. The simulation scenario contains two group-targets and three independent point-targets. The first group contains five targets and the second group contains three targets, each group-target having an arbitrarily dynamic shape. The distance between the members of the group-target vary along the simulation frames, where the maximum distance between a member of the group-target and its nearest neighbour in the same group is $d = 60m$, while the closest an out-group target gets to another target is $67m$. Looking at this from different perspectives, we can say; to correctly identify the group-targets in this scenario, it is required to separate targets $7m$ apart from each other. In the simulation, the target Cartesian position is assumed to be measured by a noisy sensor with a white Gaussian noise characterised by a covariance matrix $\mathbf{R} = \text{diag}[12^2, 12^2]$ and scan rate 400×10^{-3} sec. Using the reported measurements

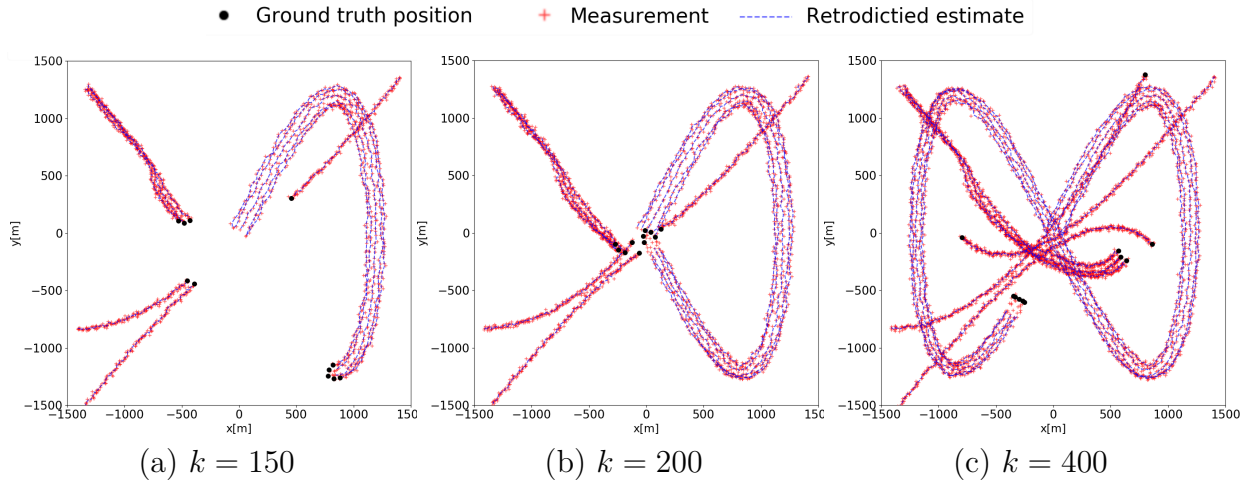


Fig. 5.4 In this simulation scenario, two group-targets and three point-targets are used. The blue dashed line shows the retrodicted states of the targets correspond to time frame $k - 1$.

and under a perfect association assumption, each target is tracked independently. Fig. 5.4 shows the simulation scenario at three different frames.

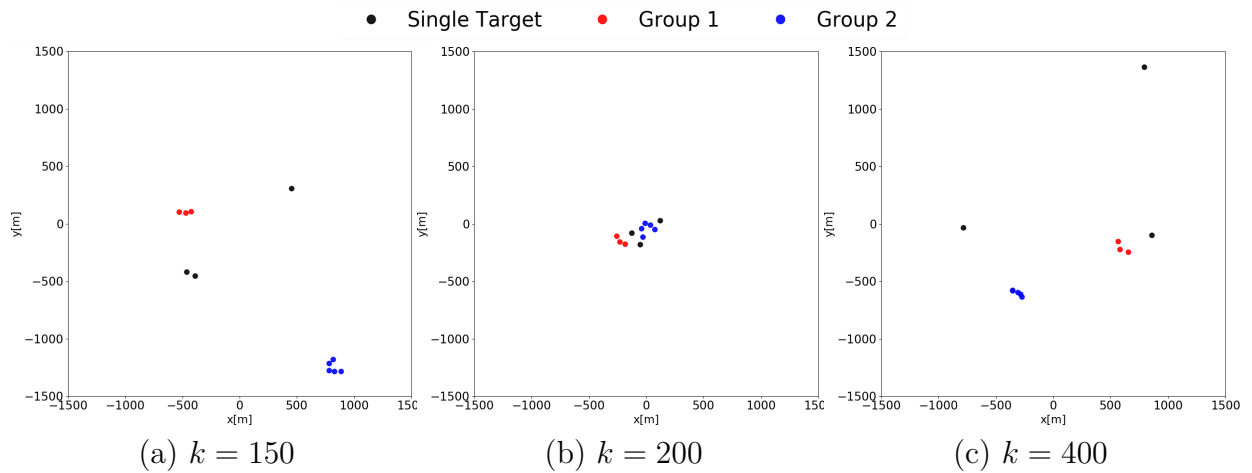


Fig. 5.5 The figure illustrates the identified group-targets and point-targets by DBSCAN using $ASDF_{n=20}$, which are coded by color at a different simulation frames.

Following that, the DBSCAN algorithm is used to identify the group-targets/clusters. The value of the DBSCAN parameter ϵ is chosen by performing a similar test shown in Fig. 5.3 but with modified distances to fit the desired constraint. Accordingly, the value of ϵ is set to the threshold which produces the best separability for the Kalman filter, $ASDF_{n=5}$, $ASDF_{n=10}$ and $ASDF_{n=20}$ as 12.5 ,40 ,74 ,143, respectively. The

other parameter $MinPts$ value is set to two, therefore, a group must be identified if two targets are travelling with a relative distance smaller than set constraint d . Then, the distance metric (5.29) is computed between each track to form a symmetric distance matrix, which, in turn, is used by the DBSCAN algorithm function. Fig. 5.10 illustrates the resulting colour-coded group by the DBSCAN algorithm based on ASDs with a window size $n = 20$.

The metric of choice to compare the performance of the DBSCAN using Kalman filter and ASDs is number of the wrong clusters. A wrong cluster is counted in the following cases;

- If in-group member is labelled as a single target.
- If in-group member is labelled as a member of different group-target.
- If a point-target is labelled as a member of a group-target.

A 1000 Monte Carlo run of the simulation is used to count the number of wrong clusters produced by the DBSCAN based on different targets. The advantage of using ASDs of a single state (Kalman filter) to improve the separability of the target's state thus, identification of the of group-target is clearly demonstrated. In Fig. 5.6, the number of the wrong clusters of a 1000 Monte Carlo run of the simulation framework based on different state densities are illustrated.

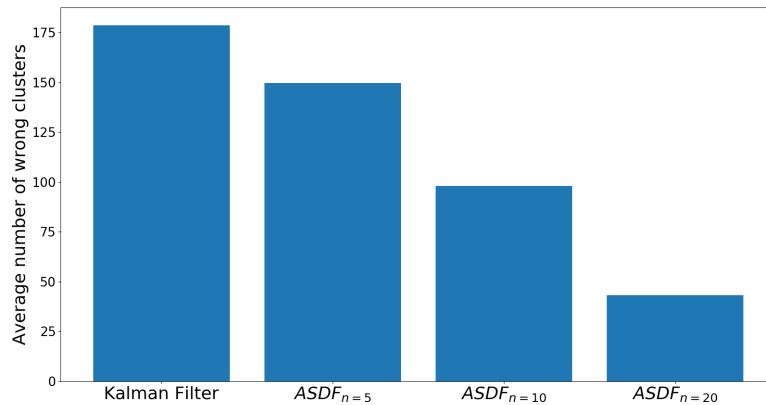


Fig. 5.6 The figure shows the average number of wrong clusters of DBSCAN with Kalman filter comparing to DBSCAN with ASDF using different time window sizes.

5.5 Automatic Identification of Coordinated Targets

The advancements in Unmanned Systems are transforming the threats that traditional defence systems are designed to tackle. One significant development is the enhancement of artificial intelligence capabilities, which enables groups of agents to perform complex collective behaviours. Consequently, equipping defence systems with threat intelligence capabilities has become essential, as identifying collective behaviours and coordinated targets can significantly enhance the system's effectiveness. In group-target tracking, the primary method to detect group targets relies on the targets being closely spaced. Therefore, spatial clustering or analysing the kinematic similarity of the targets is sufficient for detection. However, this assumption does not hold for coordinated targets. Swarming is a long-standing military tactic where multiple units converge to attack a target from single or multiple axes in a deliberately structured and coordinated manner. Such coordinated targets follow a set of rules that each group member adheres to in order to achieve a shared goal. Members of a coordinated group can spread out to attack a target from different angles or be widely separated to search a large area. Thus, the methods used for detecting group targets cannot be directly applied to coordinated targets.

In this work, members of a coordinated group are assumed to move toward a common point (converge). An approach to detect coordinated targets of this type is proposed. The proposed approach utilised the geometry of targets' velocity vectors to infer the degree of correlation over a sliding time window. Based on that, coordinated/correlated targets appear statistically closer, therefore, target pairs which move toward a common point have a higher likelihood of belonging to the same group compared to other targets. The obtained likelihood serves the role of an observation model which is used in a Bayesian framework to calculate the most likely group structure formed by the targets. Each target is assumed a point-target, in other words, at most one measurement could stem from the target. The target's kinematic state vector is accumulated over a time window and an Accumulated State Density filter (ASDF) is used to estimate the accumulated state which fully describes the correlation between the state estimates at different time instants. It has been shown that the use of an Accumulated State Densities (ASDs) improves the separability between the group's members thus the clustering[2]. In this section, we discuss a probabilistic approach to identify coordinated targets of certain types. A scoring function that utilises ASDs over a sliding time window is proposed to compute the likelihood of a pair of targets being coordinated.

A simulation scenario including different types of coordination is used to test the performance of this proposed method.

5.5.1 Problem Statement

Let $\mathcal{X}_t = \{\mathbf{x}_{k:n}^1, \mathbf{x}_{k:n}^2, \dots, \mathbf{x}_{k:n}^m\}$ be a set of accumulated states related to m resolved targets. The vector $\mathbf{x}_{k:n}^i = [\mathbf{x}_k^i, \dots, \mathbf{x}_n^i]^\top$ contains all information on the i^{th} target accumulated states over a time window of $t_k, t_k - 1, \dots, t_n$ of length $k - n + 1$. The goal of this work is to detect coordinated targets and assign a unique label to the set of targets with a common objective. In other words, calculating at each time instant t the most likely group structure G_t formed by the m targets. An example originally presented in [34] of a one possible group structure formed by 5 targets is $G_t = [1, 1, 2, 2, 3]^T$, which means that targets 1 and 2 are in group 1, targets 3 and 4 are in group 2, and target 5 is in group 3 Fig.5.7. The first target is always labelled

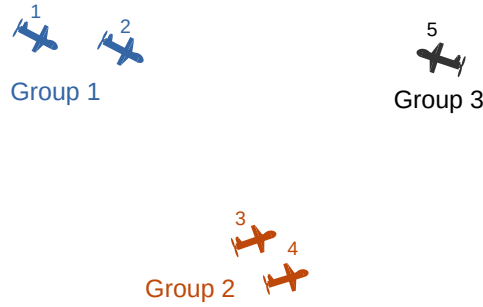


Fig. 5.7 Illustration of the group structure $G_t = [1, 1, 2, 2, 3]^T$ formed by 5 targets.

as group 1. A single target is also considered a group. Hence, the maximum number of possible groups denoted by r is also 5. In this way, the number of possible group structures for 5 targets is given by the 5th Bell number $B_5 = 52$. This can also be viewed as the number of distinct possible partitions which can be formed by a set of m targets. This is equivalent to the number of various possible group structures.

In the previous section, we discussed an approach that demonstrates how the separability between group members can be improved, thereby enhancing clustering by utilising ASDs. The focus of this section is to detect coordinated groups of the types *line shape* and *star shape*, where the assumptions that group members are closely spaced and share common kinematics are not required.

5.5.2 Probability of a group structure

Consider the set of m targets' accumulated state $\mathcal{X}_t = \{\mathbf{x}_{k:n}^1, \mathbf{x}_{k:n}^2, \dots, \mathbf{x}_{k:n}^m\}$ and their corresponding covariance $\mathbf{P}_t = \{\mathbf{P}_{k:n}^1, \mathbf{P}_{k:n}^2, \dots, \mathbf{P}_{k:n}^m\}$, it is desired to calculate the probability $p(G_t|\mathcal{X}_t)$ describing the relative likelihood of a group structure G_t given the information contained in the set $\mathcal{X}_t$. This quantity can be computed by applying Bayes theorem,

$$\begin{aligned} p(G_t|\mathcal{X}_t) &= c p(G_t) p(\mathcal{X}_t|G_t) \\ &= c p(G_t) p(\mathbf{x}_{k:n}^1, \mathbf{x}_{k:n}^2, \dots, \mathbf{x}_{k:n}^m|G_t), \end{aligned} \quad (5.30)$$

by dropping the normalisation constant c , the posterior can be described up to proportionality, that is,

$$p(G_t|\mathcal{X}_t) \propto p(G_t) p(\mathbf{x}_{k:n}^1, \mathbf{x}_{k:n}^2, \dots, \mathbf{x}_{k:n}^m|G_t). \quad (5.31)$$

To compute the posterior, it is required to have knowledge about the joint probability $P(\mathbf{x}_{k:n}^1, \mathbf{x}_{k:n}^2, \dots, \mathbf{x}_{k:n}^m|G_t)$, this probability describes how likely the current targets' accumulated states given a group structure $G_t = h$. Thus, $P(\mathbf{x}_{k:n}^1, \mathbf{x}_{k:n}^2, \dots, \mathbf{x}_{k:n}^m|G_t = h)$ serves the role of an observation model. This joint probability can be factorised using the chain rule for probability as

$$\begin{aligned} p(\mathbf{x}_{k:n}^1, \mathbf{x}_{k:n}^2, \dots, \mathbf{x}_{k:n}^m|G_t = h) &= p(\mathbf{x}_{k:n}^1|h) p(\mathbf{x}_{k:n}^2|\mathbf{x}_{k:n}^1, h) \\ &\quad \dots p(\mathbf{x}_{k:n}^m|\mathbf{x}_{k:n}^1, \dots, \mathbf{x}_{k:n}^{m-1}, h) \end{aligned} \quad (5.32)$$

In order to compute a numerical quantity for the probability above, we need to define a general expression for the conditional probability $p(\mathbf{x}_{k:n}^i|\mathbf{x}_{k:n}^1, \dots, \mathbf{x}_{k:n}^j, \dots, \mathbf{x}_{k:n}^q, h) : q \leq m-1, i \neq j$.

Suppose we have a scoring function $f(\mathbf{x}_{k:n}^i, \mathbf{x}_{k:n}^j) \in [0, 1]$ that measures the likelihood for a pair of targets i and j to be in the same group. With this, we define the conditional probability

$$\begin{aligned} p(\mathbf{x}_{k:n}^i|\mathbf{x}_{k:n}^1, \dots, \mathbf{x}_{k:n}^j, \dots, \mathbf{x}_{k:n}^q, h) &= \prod_{j \in g}^q f(\mathbf{x}_{k:n}^i, \mathbf{x}_{k:n}^j) \\ &\quad \prod_{j \notin g}^q [1 - f(\mathbf{x}_{k:n}^i, \mathbf{x}_{k:n}^j)] \end{aligned} \quad (5.33)$$

where g denotes the group label of target $\mathbf{x}_{k:n}^i$. By substituting the definition (5.33) in (5.32), one can show that the resulting factorisation will include all possible combinations of pairs in the set $\mathcal{X}_t$. It should be mentioned that in case we have a single variable as

in the term $p(\mathbf{x}^i|h)$, it is not possible to infer the probability of the group structure, for this reason, the term can be dropped in (5.32), or set to equal 0.5, which represents a complete ignorance.

Measure of Coordination

We wish to define a scoring function $f(\mathbf{x}_{k:n}^i, \mathbf{x}_{k:n}^j) \in [0, 1]$ that measures the degree of similarity/coordination between a pair of targets i and j . A feasible definition of $f(\mathbf{x}_{k:n}^i, \mathbf{x}_{k:n}^j)$ that serves this purpose is:

$$f(\mathbf{x}_{k:n}^i, \mathbf{x}_{k:n}^j) = e^{-\alpha(\mathbf{x}_{k:n}^i - \mathbf{x}_{k:n}^j)^\top \mathbf{P}_{k:n}^\Delta^{-1} (\mathbf{x}_{k:n}^i - \mathbf{x}_{k:n}^j)} \quad (5.34)$$

where α is a constant related to the desired distance in which pairs are allowed to be in the same group. The proposed statistical score is based on the normalised difference between the two estimates

$$\Delta_{k:n}^{i,j} = \mathbf{x}_{k:n}^i - \mathbf{x}_{k:n}^j \quad (5.35)$$

by the covariance $\mathbf{P}_{k:n}^\Delta$, which is written after dropping the time index for brevity as:

$$\begin{aligned} \mathbf{P}_{k:n}^\Delta &= \text{cov}(\mathbf{x}^i - \mathbf{x}^j) \\ &= \left[\left((\mathbf{x}^i - \mathbf{x}^j) - [\mathbf{x}^i - \mathbf{x}^j] \right) \left((\mathbf{x}^i - \mathbf{x}^j) - [\mathbf{x}^i - \mathbf{x}^j] \right)^\top \right] \\ &= \mathbf{P}^i + \mathbf{P}^j - 2\mathbf{P}^{ij} \end{aligned} \quad (5.36)$$

where

$$\mathbf{P}^{ij} = \left[(\mathbf{x}^i - [\mathbf{x}^i])(\mathbf{x}^j - [\mathbf{x}^j])^\top \right] \quad (5.37)$$

is the cross-covariance matrix which reflects the correlation between the two targets' estimates. Usually, it is hard to compute the cross-covariance term $\mathbf{P}^{ij}$. However, the cross-covariance is related to the degree of coordination, in other words, the correlation. The relation between the cross-covariance and the correlation is described in the definition of the correlation coefficient $\rho^{i,j}$ between two random variables $\mathbf{x}^i$ and $\mathbf{x}^j$ as follows:

$$\rho^{i,j} = \mathbf{P}^{ij}(\mathbf{P}^i \mathbf{P}^j)^{-\frac{1}{2}}, \quad (5.38)$$

where the term $(*)^{-\frac{1}{2}}$ is computed using Cholesky decomposition. By rearranging (5.38), the cross-covariance can be computed as:

$$\mathbf{P}^{ij} = \rho^{i,j} (\mathbf{P}^i \mathbf{P}^j)^{\frac{1}{2}}. \quad (5.39)$$

All information on the target states accumulated over the selected time window is contained in the ASD. If one analyses the ASD from the perspective of time series, one can utilise the available knowledge about the targets' accumulated states in the selected time window to compute correlation coefficient $\rho^{i,j}$. We define coordination as when several targets converge to attack a certain point from single or multiple axes. This implies that coordinated targets must be heading toward a common point in the coordinate system. Thus, targets that move toward a common point over time would have a higher correlation, in other words, are judged more likely to form a group. This definition can be understood as the geometry between velocity vectors related to the pair of targets, as illustrated in Fig.5.8. Therefore, to distinguish between converging and diverging pair, the "radial velocity" or "line-of-sight velocity" can be used. The "radial velocity" is equivalent to the vector projection of the target's relative velocity onto the relative direction connecting the two targets. The velocity component resulting from projecting target i velocity vector $\mathbf{V}^i$ onto the vector $\bar{\mathbf{r}}$ connecting target i and j , can be computed as:

$$v^{r\angle i} = \frac{\mathbf{V}^i \cdot \bar{\mathbf{r}}}{\|\bar{\mathbf{r}}\|} \quad (5.40)$$

where the operator $\cdot$ denotes a dot product, and $\|\dots\|$ is the vector norm. Similarly, we compute the component related to target j ,

$$v^{r\angle j} = \frac{\mathbf{V}^j \cdot \bar{\mathbf{r}}}{\|\bar{\mathbf{r}}\|} \quad (5.41)$$

As the radial velocity is relative, it remains to subtract the components resulting from the projection

$$\begin{aligned} V_{rad}^{i,j} &= v^{r\angle i} - v^{r\angle j} \\ &= (\mathbf{V}^i - \mathbf{V}^j) \cdot \frac{\bar{\mathbf{r}}}{\|\bar{\mathbf{r}}\|} \\ &= \frac{\|\mathbf{V}^i - \mathbf{V}^j\| \|\bar{\mathbf{r}}\| \cos(\theta)}{\|\bar{\mathbf{r}}\|} \\ &= \|\mathbf{V}^i - \mathbf{V}^j\| \cos(\theta) \end{aligned} \quad (5.42)$$

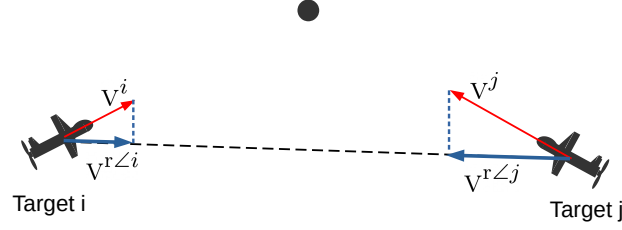


Fig. 5.8 The heading of targets i and j shows that both targets are moving toward a common point. The radial velocity components resulting from projecting the velocity vector on the vector connecting i and j are shown in blue.

The computed radial velocity $v_{rad}^{i,j}$ tells us whether the targets are converging or not, however, we are interested in the degree of correlation. The degree of correlation can be directly obtained from $\cos(\theta)$, which represents the angle between the relative velocity vector and the vector $\|v^i - v^j\|$ vector on the line-of-sight between the pair of targets $\bar{r}$. For example, if the relative velocity and line-of-sight are orthogonal ($\theta = \frac{\pi}{2}$), then $\cos(\theta) = 0$, i.e, the pair are not correlated which corresponds to a pair moving parallel to each other at a similar speed. To further convince the reader that $\cos(\theta)$ can be interpreted as a correlation coefficient, is that, both ρ^{ij} and $\cos(\theta)$ have numerical value in the range $\in [-1, 1]$. With that we can compute correlation coefficient ρ^{ij} as:

$$\begin{aligned} \rho^{ij} &= \cos(\theta) \\ &= \frac{(v^i - v^j) \cdot \bar{r}}{\|v^i - v^j\| \|\bar{r}\|} \end{aligned} \quad (5.43)$$

By substituting the result above in (5.39), the term of the cross-covariance become:

$$\mathbf{P}^{ij} = \frac{(v^i - v^j) \cdot \bar{r}}{\|v^i - v^j\| \|\bar{r}\|} (\mathbf{P}^i \mathbf{P}^j)^{\frac{1}{2}}. \quad (5.44)$$

the effects of the cross-covariance calculated via (5.44) on the pair's likelihood (5.34), is illustrated in Fig.5.9.

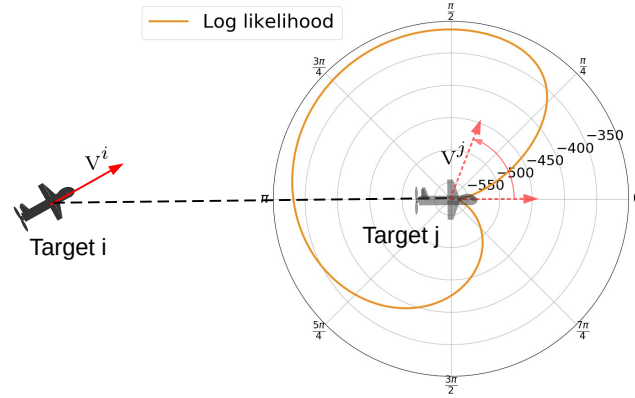


Fig. 5.9 Target i and j are spread by a fixed distance. The heading of target i is also fixed to about 45 degrees, while target j 's heading rotates along the circle. Also, target i 's velocity is greater than target j 's velocity. The orange curve shows the different log-likelihood of targets i and j to belong to the same group for each different heading of j . The log-likelihood is obtained by taking the log of $f(\mathbf{x}_{k:n}^i, \mathbf{x}_{k:n}^j)$ defined in (5.34).

5.5.3 Simulation and Results

The simulated scenario designed to test the proposed approach includes eight targets. The targets' dynamics are constructed to form three groups of the distinguished coordination types, as follows: three targets form a closely spaced group, three targets form a star shape, and the last 2 targets form a line shape. Targets' positions are measured with additive Gaussian zero mean noise, where the covariance for all instants of time is given by $\mathbf{R} = \text{diag}[12^2, 12^2]$. The targets are simulated to move according to the constant turn rate model with various velocities depending on the group, the maximum absolute velocity used is 27m/s . In the absence of clutter measurements and under the perfect sensor-data-to-track association, accumulated targets' states over a time window of length 6 are estimated using ASD filter. The parameter of the scoring function (5.34) is targets pair's ASDs, where the kinematic state vector of a single instance of time $\mathbf{x}_t = [x, y, v, v]^\top$, this kinematic state is a reduced state derived from the kinematic state estimated via the ASD filter. Selected snapshots of the simulation at different frames are shown in Fig. 5.10. The snapshots show colour-coded targets' positions at different instances of the simulation, the colour code represents the different groups of the most likely group structure computed by the proposed approach. It can be seen that targets that appear to move toward a common point are clustered as a coordinated group. Targets that are relatively closely spaced (from Euclidean distance perspective) from a group are labelled as independent once they start to steer away from the rest of the group, Fig. 5.10 (b).

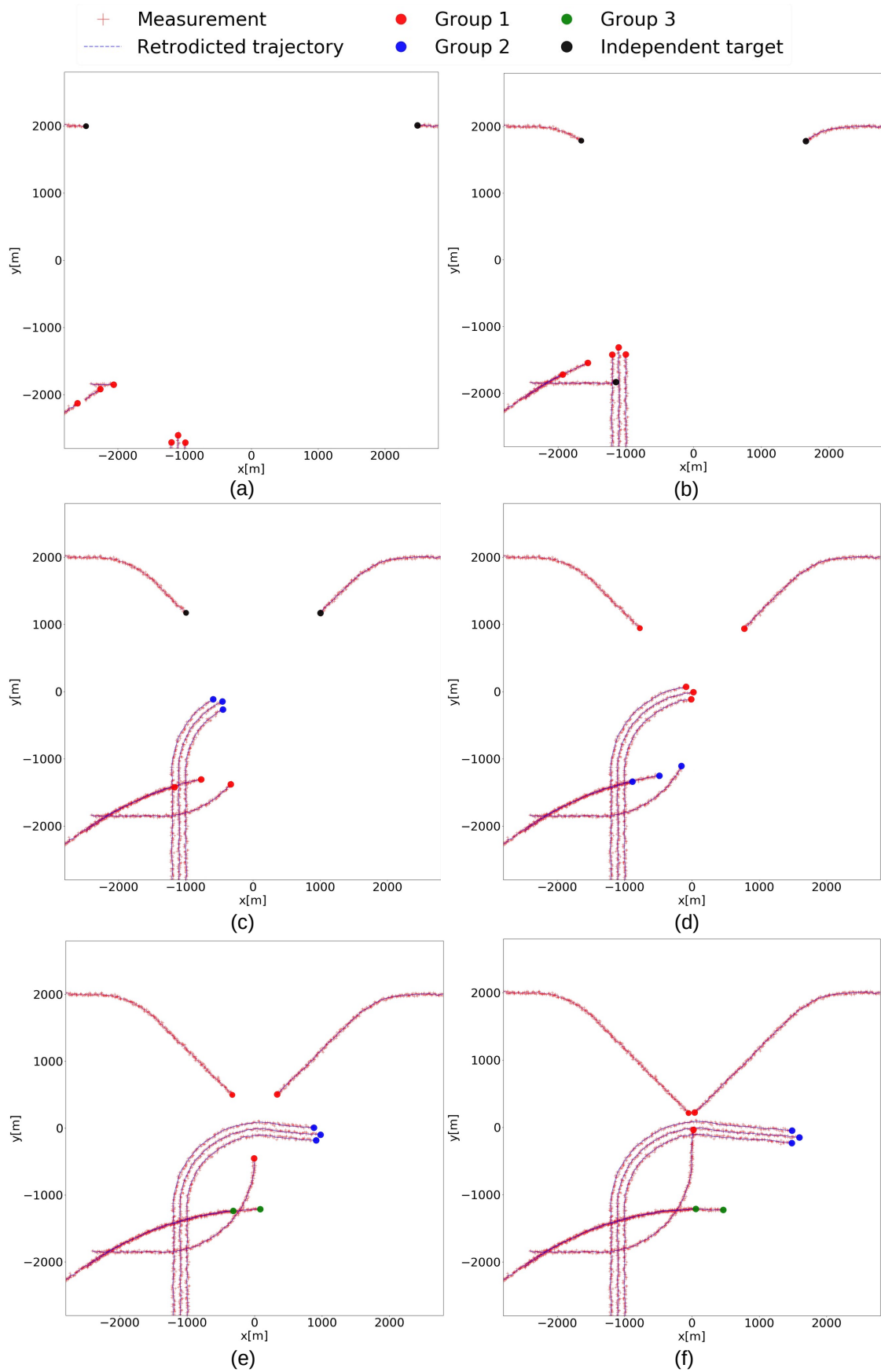


Fig. 5.10 Illustrates the most likely group structure at different frames which are represented by colour-coded groups and independent targets.

5.6 Conclusion

In this chapter, we discussed the problem of coordinated target tracking in two sections. The first section focused on closely spaced group targets, while the second section relaxed the closely spaced assumption. We demonstrated in the first section that the separability of tracks could be significantly enhanced by using ASDs instead of a single state, as is typically done with a Kalman filter. This is especially important in applications where high accuracy in the distance between tracks is required. It is worth noting that the level of achievable separability is significantly influenced by factors such as sensor measurement noise, process noise, and the types of manoeuvres performed by the targets.

In the second section, we addressed the problem of detecting collaborative behaviour, specifically the convergence of multiple units to achieve a common objective. We formulated this problem using a Bayesian approach to identify the most likely group structure. The observation model in the Bayesian formula was obtained by factorising the joint probability density of the targets, given the group structure. To compute a numerical value for this factorisation, we proposed a scoring function that measures the likelihood of a pair of targets belonging to the same collaborative group. Additionally, we introduced a measure of the correlation coefficient between pairs of accumulated states over a time window, demonstrating that pairs moving toward a common point become statistically closer than pairs steering away from each other.

One challenge is that the number of possible group structures corresponds to the number of ways to partition a set of m targets, which is a Bell number. This number can be significantly reduced by introducing a gating mechanism that allows unlikely groupings to be ignored. We observed that increasing the sliding window of the ASDs improved the performance of this approach. However, it is crucial to find the optimal trade-off between computational cost and performance when selecting the ASDs' sliding window.

Chapter 6

Dissertation Conclusion

The work presented in this dissertation primarily focused on extended and coordinated target tracking. Since group target tracking problems can fall under both categories, and given that the work related to coordinated target tracking only considered the aspect of identifying coordination, group target tracking was addressed in the latter context. One can utilise approaches for the simultaneous estimation of shape and kinematics (extended target) and combine them with automatic detection methods for group targets. In this work, we assumed that the shape of the group is independent; however, it is worth investigating in the future whether it is possible to infer more information about group behaviour based on the estimated shape.

As the author writes the final chapter of his PhD work, he can argue that research in ETT has reached a very advanced stage. Fast, robust, and accurate algorithms are now available for deployment in real-time products; however, there is still work to be done in this area. On the other hand, the field of coordinated target tracking is relatively unexplored. Given the significant interest and advancements in multi-agent systems, the author anticipates an increasing need for systems capable of identifying collective and collaborative behaviour¹. Upon identifying coordinated behaviour, we need a system to be able to take appropriate actions, manage sensors and effectors, refine the situation, and gather threat intelligence to minimise the impact of potential threats through multi-coordinated responses. This is the final concept the author aimed to highlight, and with this, we can conclude our discussion and the author's PhD work.

¹If computer scientists are able to provide effective solutions in this area, biologists would be particularly grateful. In fact, following the publication of a paper on automatic identification of coordinated targets, the author was approached by biologists seeking assistance in developing methods to identify specific behaviours in wild animals (As a computer scientist, your life can be both fun and exciting, which may be different from what many women might assume.)

The conclusions regarding the methods and approaches proposed in the dissertation are not discussed here but are instead presented at the end of each chapter or section to avoid repetition of the same information.

References

- [1] Alqaderi, H., Govaers, F., and Koch, W. (2021). Symmetric star-convex shape tracking with wishart filter. In *2021 IEEE 24th International Conference on Information Fusion (FUSION)*, pages 1–8.
- [2] Alqaderi, H., Govaers, F., and Koch, W. (2022a). Accumulated state densities filter for better separability of group-targets. In *2022 25th International Conference on Information Fusion (FUSION)*, pages 1–7.
- [3] Alqaderi, H., Govaers, F., and Koch, W. (2022b). Bayesian wishart filter for random shape tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 58(3):1941–1952.
- [4] Alqaderi, H., Govaers, F., and Schulz, R. (2019). Spacial elliptical model for extended target tracking using laser measurements. In *2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pages 1–6.
- [5] Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. John Wiley Sons.
- [6] Bar-Shalom, Y., Li, X. R., and Kirubarajan, T. (2004). *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons.
- [7] Baum, M., Faion, F., and Hanebeck, U. D. (2012). Modeling the target extent with multiplicative noise. In *2012 15th International Conference on Information Fusion*, pages 2406–2412.
- [8] Baum, M. and Hanebeck, U. D. (2011). Shape tracking of extended objects and group targets with star-convex rhms. In *14th International Conference on Information Fusion*, pages 1–8. IEEE.
- [9] Baum, M. and Hanebeck, U. D. (2014). Extended object tracking with random hypersurface models. *IEEE Transactions on Aerospace and Electronic systems*, 50(1):149–159.
- [10] Braca, P., Aubry, A., Millefiori, L. M., De Maio, A., and Marano, S. (2020). Multi-class random matrix filtering for adaptive learning. *IEEE Transactions on Signal Processing*, 68:359–373.
- [11] Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). Density-based spatial clustering of applications with noise. In *Int. Conf. Knowledge Discovery and Data Mining*, volume 240, page 6.

- [12] Fackler, M. and Aktuar, D. (2009). Panjer class united—one formula for the poisson, binomial, and negative binomial distribution. In *ASTIN colloquium*.
- [13] Faion, F., Baum, M., and Hanebeck, U. D. (2012). Tracking 3D Shapes in Noisy Point Clouds with Random Hypersurface Models. In *2012 15th International Conference on Information Fusion*, pages 2230–2235. IEEE.
- [14] Feldmann, M., Fränken, D., and Koch, W. (2011). Tracking of extended objects and group targets using random matrices. *IEEE Transactions on Signal Processing*, 59(4):1409–1420.
- [15] García-Fernández, F., Svensson, L., and Moreland, M. R. (2020). Multiple target tracking based on sets of trajectories. *IEEE Transactions on Aerospace and Electronic Systems*, 56(3):1685–1707.
- [16] Govaers, F. and Alqaderi, H. (2020). A gamma filter for positive parameter estimation. In *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 40–45.
- [17] Govaers, F. and Koch, W. (2019). State and trajectory estimation using accumulated state densities. *Perspectives on Information Fusion*, 2(1):20–27.
- [18] Granström, K. and Baum, M. (2023). A tutorial on multiple extended object tracking. *Authorea Preprints*.
- [19] Granstrom, K., Baum, M., and Reuter, S. (December, 2017). Extended Object Tracking: Introduction, Overview, and Applications. *Journal of Advances in Information Fusion*, 12(2):139–174.
- [20] Granström, K., Lundquist, C., and Orguner, U. (2011). Tracking rectangular and elliptical extended targets using laser measurements. In *14th International Conference on Information Fusion*, pages 1–8. IEEE.
- [21] Granström, K., Reuter, S., Meissner, D., and Scheel, A. (2014). A multiple model phd approach to tracking of cars under an assumed rectangular shape. In *17th International Conference on Information Fusion (FUSION)*, pages 1–8.
- [22] Gunnarsson, J., Svensson, L., Danielsson, L., and Bengtsson, F. (2007). Tracking vehicles using radar detections. In *2007 IEEE Intelligent Vehicles Symposium*, pages 296–302.
- [23] Hammarstrand, L., Svensson, L., Sandblom, F., and Sorstedt, J. (2012). Extended object tracking using a radar resolution model. *IEEE Transactions on Aerospace and Electronic Systems*, 48(3):2371–2386.
- [24] Kaulbersch, H., Honer, J., and Baum, M. (2018). A cartesian b-spline vehicle model for extended object tracking. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 1–5. IEEE.
- [25] Koch, J. W. (2008a). Bayesian approach to extended object and cluster tracking using random matrices. *IEEE Transactions on Aerospace and Electronic Systems*, 44(3):1042–1059.

- [26] Koch, W. (2008b). Bayesian Approach to Extended Object and Cluster Tracking using Random Matrices. *IEEE Transactions on Aerospace and Electronic Systems*, 44(3):1042–1059.
- [27] Koch, W. (2009). On accumulated state densities with applications to out-of-sequence measurement processing. In *2009 12th International Conference on Information Fusion*, pages 2201–2208.
- [28] Koch, W. (2014). *Tracking and Sensor Data Fusion: Methodological Framework and Selected Applications*. Springer.
- [29] Koch, W., Govaers, F., and Charlish, A. (2013). An exact solution to track-to-track fusion using accumulated state densities. In *2013 Workshop on Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pages 1–5.
- [30] Liu, W., Zhu, S., and Wen, C. (2016). Multiple resolvable group target estimation using graph theory and the multi-bernoulli filter. In *2016 International Conference on Control, Automation and Information Sciences (ICCAIS)*, pages 21–26.
- [31] Mahler, R. (2003a). Multitarget bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4):1152–1178.
- [32] Mahler, R. P. (2003b). Multitarget bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic systems*, 39(4):1152–1178.
- [33] Mathai, A. and Moschopoulos, P. (1991). On a multivariate gamma. *Journal of Multivariate Analysis*, 39(1):135–153.
- [34] Pang, S. K., Li, J., and Godsill, S. J. (2011). Detection and tracking of coordinated groups. *IEEE Transactions on Aerospace and Electronic Systems*, 47(1):472–502.
- [35] Rao, C. R., Rao, C. R., Statistiker, M., Rao, C. R., and Rao, C. R. (1973). *Linear Statistical Inference and its Applications*. Wiley New York.
- [36] Rauch, H. E., Tung, F., and Striebel, C. T. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA journal*, 3(8):1445–1450.
- [37] Rong Li, X. and Jilkov, V. (2003). Survey of maneuvering target tracking. part i. dynamic models. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4):1333–1364.
- [38] Ruud, K. A., Brekke, E. F., and Eidsvik, J. (2018). Lidar extended object tracking of a maritime vessel using an ellipsoidal contour model. In *2018 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pages 1–6. IEEE.
- [39] Schlangen, I. C. (2017). *Multi-object filtering with second-order moment statistics*. PhD thesis, Heriot-Watt University.
- [40] Schleiss, M. (2019). Translating aerial images into street-map-like representations for visual self-localization of uavs. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*.

- [41] Schuster, M. and Reuter, J. (2015). Target Tracking in Marine Environment using Automotive Radar and Laser Range Sensor. In *2015 20th International Conference on Methods and Models in Automation and Robotics (MMAR)*, pages 965–970. IEEE.
- [42] Straka, O., Duník, J., and Šimandl, M. (2012). Truncation nonlinear filters for state estimation with nonlinear inequality constraints. *Automatica*, 48(2):273–286.
- [43] Streit, R. L. and Stone, L. D. (2008). Bayes derivation of multitarget intensity filters. In *2008 11th International Conference on Information Fusion*, pages 1–8.
- [44] Sun, L., Lan, J., and Li, X. R. (2012). Extended Target Tracking using Star-Convex Model with Nonlinear Inequality Constraints. In *Proceedings of the 31st Chinese Control Conference*, pages 3869–3874.
- [45] Tuncer, B. and Özkan, E. (2021). Random matrix based extended target tracking with orientation: A new model and inference. *IEEE Transactions on Signal Processing*, 69:1910–1923.
- [46] Viraswami, K. (1991). *On Multivariate Gamma Distributions*. McGill University.
- [47] Wahlström, N. and Özkan, E. (2015). Extended Target Tracking using Gaussian Processes. *IEEE Transactions on Signal Processing*, 63(16):4165–4178.
- [48] Williams, C. K. and Rasmussen, C. E. (2006). *Gaussian Processes for Machine Learning*. MIT press Cambridge, MA.
- [49] Williams, S. M. (2018). Swarm weapons: Demonstrating a swarm intelligent algorithm for parallel attack. Technical report, US Army School for Advanced Military Studies Fort Leavenworth United States.
- [50] Xu, L., Li, X. R., Liang, Y., and Duan, Z. (2017). Constrained dynamic systems: Generalized modeling and state estimation. *IEEE Transactions on Aerospace and Electronic Systems*, 53(5):2594–2609.
- [51] Yang, J.-l., Li, P., and Ge, H.-w. (2014). Extended target shape estimation by fitting b-spline curve. *Journal of Applied Mathematics*, 2014.

Appendix A

A.1 List of Publications

The papers published in relation to this dissertation are listed in the following two lists:

Publications Related to Extended Target Tracking

- H. Alqaderi, F. Govaers and R. Schulz, "Spacial Elliptical Model for Extended Target Tracking Using Laser Measurements," 2019 Sensor Data Fusion: Trends, Solutions, Applications (SDF), Bonn, Germany, 2019, pp. 1-6, doi: 10.1109/SDF.2019.8916634.
- F. Govaers and H. Alqaderi, "A Gamma Filter for Positive Parameter Estimation," 2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), Karlsruhe, Germany, 2020, pp. 40-45, doi: 10.1109/MFI49285.2020.9235265.
- H. Alqaderi, F. Govaers and W. Koch, "Symmetric Star-convex Shape Tracking With Wishart Filter," 2021 IEEE 24th International Conference on Information Fusion (FUSION), Sun City, South Africa, 2021, pp. 1-8, doi: 10.23919/FUSION49465.2021.9626838.
- H. Alqaderi, F. Govaers and W. Koch, "Bayesian Wishart Filter for Random Shape Tracking," in IEEE Transactions on Aerospace and Electronic Systems, vol. 58, no. 3, pp. 1941-1952, June 2022, doi: 10.1109/TAES.2021.3131965.

Publications Related to Coordinated Target Tracking

- H. Alqaderi, F. Govaers and W. Koch, "Accumulated State Densities Filter for Better Separability of Group-Targets," 2022 25th International Conference on In-

formation Fusion (FUSION), Linköping, Sweden, 2022, pp. 1-7, doi: 10.23919/FUSION49751.2022.9841279.

- H. Alqaderi, F. Govaers and W. Koch, "Automatic Identification of Coordinated Targets," 2023 26th International Conference on Information Fusion (FUSION), Charleston, SC, USA, 2023, pp. 1-7, doi: 10.23919/FUSION52260.2023.10224124.

Appendix B

B.1 Multiplication of Wishart density functions

Giving two Wishart distributed densities of the same dimensions d ; $\mathcal{W}(\mathbf{X}; n_z, \mathbf{C}_z)$ and $\mathcal{W}(\mathbf{X}; n_0, \mathbf{C}_0)$, which describe the measurement likelihood and the prior, respectively. The product of the two densities is calculated as follows:

$$\begin{aligned}
\mathcal{W}(\mathbf{X}; n, \mathbf{C}) &\propto \mathcal{W}(\mathbf{X}; n_z, \mathbf{C}_z) \mathcal{W}(\mathbf{X}; n_0, \mathbf{C}_0) \\
&\propto |\mathbf{X}|^{\frac{n_z-d-1}{2}} e^{-\frac{1}{2} \text{tr}(\mathbf{C}_z^{-1} \mathbf{X})} |\mathbf{X}|^{\frac{n_0-d-1}{2}} e^{-\frac{1}{2} \text{tr}(\mathbf{C}_0^{-1} \mathbf{X})} \\
&\propto |\mathbf{X}|^{\frac{n_z+n_0-d-1-d-1}{2}} e^{-\frac{1}{2} \text{tr}((\mathbf{C}_z^{-1} + \mathbf{C}_0^{-1}) \mathbf{X})} \\
&\propto |\mathbf{X}|^{\frac{n_z+n_0-d-1-d-1}{2}} e^{-\frac{1}{2} \text{tr}((\mathbf{C}_0^{-1} (\mathbf{C}_z + \mathbf{C}_0) \mathbf{C}_z^{-1}) \mathbf{X})} \\
&\propto |\mathbf{X}|^{\frac{n_z+n_0-d-1-d-1}{2}} e^{-\frac{1}{2} \text{tr}((\mathbf{C}_0 (\mathbf{C}_z + \mathbf{C}_0)^{-1} \mathbf{C}_z)^{-1} \mathbf{X})} \\
&\propto |\mathbf{X}|^{\frac{n-d-1}{2}} e^{-\frac{1}{2} \text{tr}(\mathbf{C}^{-1} \mathbf{X})}
\end{aligned} \tag{B.1}$$

B.2 Unbiased Wishart filter

We can begin the unbiasedness test by substituting the values of the Wishart posterior parameters (4.78) and (4.79). For brevity, the sub-index " $k-1$ " is dropped in the rest of this section. Instead, we write " 0 " to refer to the prior parameters.

$$\begin{aligned}
\mathbb{E}[\mathbf{X} - n\mathbf{C}] &= 0 \\
&= \mathbb{E}[\mathbf{X} - (n_z + n_0 - d - 1) \mathbf{C}_0 (\mathbf{C}_z + \mathbf{C}_0)^{-1} \mathbf{C}_z] \\
&= \mathbb{E} \left[\mathbf{X} - (n_z + n_0 - d - 1) \right. \\
&\quad \left. \frac{\sum_{j=1}^{n_0} \mathbf{z}_0^j \mathbf{z}_0^{j\top}}{n_0} \left(\frac{\sum_{j=1}^{n_z} \mathbf{z}_k^j \mathbf{z}_k^{j\top}}{n_z} + \frac{\sum_{j=1}^{n_0} \mathbf{z}_0^j \mathbf{z}_0^{j\top}}{n_0} \right)^{-1} \frac{\sum_{j=1}^{n_z} \mathbf{z}_k^j \mathbf{z}_k^{j\top}}{n_z} \right]
\end{aligned} \tag{B.2}$$

As can be seen in (B.2), the inverse is hard to calculate. To simplify the problem, the unbiasedness check can be completed by fusing the same measurement two times. In other words, the prior will be the same as the measurement likelihood. Therefore we can continue (B.2) as:

$$\begin{aligned}
&= \mathbb{E} \left[\mathbf{X} - (2n_z - d - 1) \right. \\
&\quad \left. \frac{\sum_{j=1}^{n_z} \mathbf{z}_k^j \mathbf{z}_k^{j\top}}{n_z} \frac{n_z}{2} \left(\sum_{j=1}^{n_z} \mathbf{z}_k^j \mathbf{z}_k^{j\top} \right)^{-1} \frac{\sum_{j=1}^{n_z} \mathbf{z}_k^j \mathbf{z}_k^{j\top}}{n_z} \right] \\
&= \mathbb{E} \left[\mathbf{X} - (2n_z - d - 1) \frac{\sum_{j=1}^{n_z} \mathbf{z}_k^j \mathbf{z}_k^{j\top}}{2n_z} \right]
\end{aligned} \tag{B.3}$$

Obviously, the terms $-d - 1$ caused the estimator to be biased. However, its effect decreases with the increase in number of the fused measurements. Nevertheless, this effect can be eliminated by dropping $-d - 1$ from the expectation, and by doing so, (B.3) becomes:

$$\begin{aligned}
\mathbb{E}[\mathbf{X} - n\mathbf{C}] &= \mathbb{E} \left[\mathbf{X} - \sum_{j=1}^{n_z} \mathbf{z}_k^j \mathbf{z}_k^{j\top} \right] \\
&= \mathbb{E} \left[\sum_{j=1}^{n_z} f(\mathbf{x}^j) f(\mathbf{x}^j)^\top - \sum_{j=1}^{n_z} (f(\mathbf{x}^j) + \mathbf{u}_k^j) (f(\mathbf{x}^j) + \mathbf{u}_k^j)^\top \right] \\
&= \mathbb{E} \left[\sum_{j=1}^{n_z} f(\mathbf{x}^j) f(\mathbf{x}^j)^\top - \sum_{j=1}^{n_z} f(\mathbf{x}^j) f(\mathbf{x}^j)^\top \right. \\
&\quad \left. - 2 \sum_{j=1}^{n_z} f(\mathbf{x}^j) \mathbf{u}_k^j - \sum_{j=1}^{n_z} \mathbf{u}_k^j \mathbf{u}_k^{j\top} \right] \\
&= 2 \sum_{j=1}^{n_z} \mathbb{E}[f(\mathbf{x}^j)] \mathbb{E}[\mathbf{u}_k^j] - \sum_{j=1}^{n_z} \mathbb{E}[\mathbf{u}_k^j \mathbf{u}_k^{j\top}] \\
&= - \sum_{j=1}^{n_z} \mathbb{E}[\mathbf{u}_k^j \mathbf{u}_k^{j\top}] \\
&= -n_z \mathbf{R}_d
\end{aligned} \tag{B.4}$$

Thus, the unbiased estimator of the Wishart density can be obtained as:

$$\hat{\mathbf{X}} = n\mathbf{C} - n_z \mathbf{R}_d \tag{B.5}$$

where $\mathbf{R}_d$ is a $d \times d$ measurements error covariance matrix, the parameters n and $\mathbf{C}$ are calculated as:

$$n = n_z + n_0 \quad (\text{B.6})$$

$$\mathbf{C} = \mathbf{C}_0(\mathbf{C}_z + \mathbf{C}_0)^{-1}\mathbf{C}_z \quad (\text{B.7})$$

B.3 On Self-Occlusion & Measurement Likelihood of Symmetric Shape

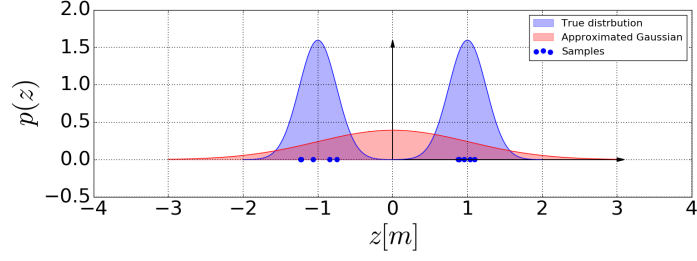
Let a symmetric segment with a radial length $r = 1 \text{ m}$ be parallel to the x-axis and centred at the origin (0,0). A measurement z stemming from one of the measurement sources is corrupted by a white Gaussian noise with variance σ_r^2 . Thus, the measurement likelihood can be described by a mixture of two Gaussian components of this form:

$$p(z|\theta^i) = \frac{1}{2}\mathcal{N}(z; r, \sigma_r^2) + \frac{1}{2}\mathcal{N}(z; -r, \sigma_r^2). \quad (\text{B.8})$$

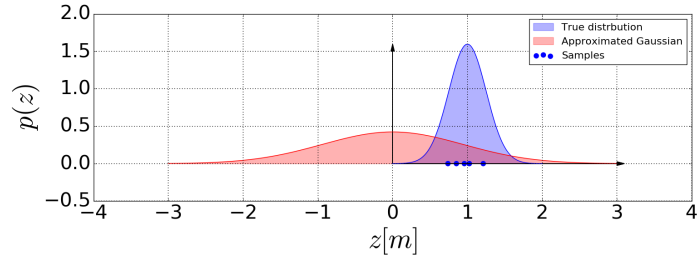
As this mixture has a zero-mean, a fair approximation could be achieved by matching its first moment to a single Gaussian density with an appropriate variance; $\mathcal{N}(z; 0, \sigma_c^2)$. Based on definition 4.3.1, the variance of the latter density is the same as the scalar matrix $\mathbf{C}$ of a uni-variate Wishart distribution. Thus, it could be estimated by the mean of the proposed Wishart filter. The resulting approximation of the mixture is shown by the red solid curve in Fig. B.1 a. In a realistic scenario, only one side of the segment is observable by a single LiDAR scan. Nevertheless, in Fig. B.1 one can see that in the uni-variate case the estimated variance appears to be similar for both scenarios. With that, we carry the assumption in our model that the measurement is distributed according to a zero-mean Gaussian, that is,

$$z \sim \mathcal{N}(z; 0, \sigma_c^2).$$

In a d-variate case, the zero-mean Gaussian assumption of the measurement likelihood needs to be handled carefully. That is because, when only one measurement source of each segment is observable, the joint distribution of the measurement likelihood $\mathbf{z} = [z^1, \dots, z^d]^\top$ not only differ from the distribution when two measurement sources are observable but also, fail to capture the true distribution. To illustrate this behaviour, Fig.B.2 compares the Wishart measurement likelihoods (4.98) and (4.99), obtained in two scenarios; with two measurement sources are observable and a single source



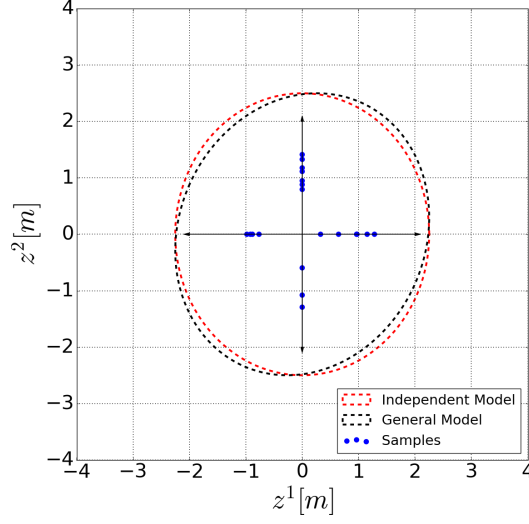
(a). Two sources are observable.



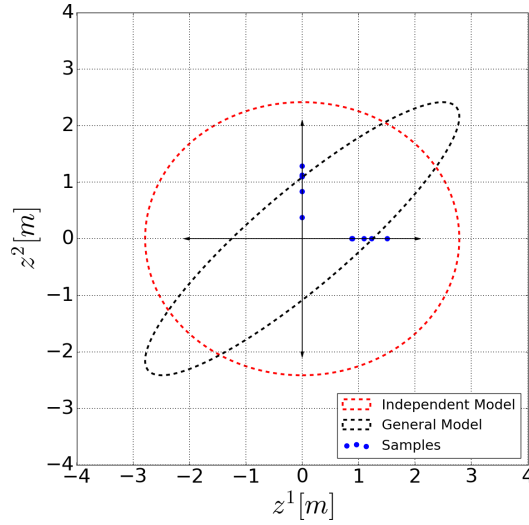
(b). A single source is observable.

Fig. B.1 This figure compares the approximation of the measurement likelihood by a uni-variate Gaussian with zero-mean in two scenarios. Scenario (a), represents the case where two ends of the segments are observable, while (b) represents the case when only one end is observable. The measurement likelihood is shown by the blue solid curve. While the red solid curve represents the approximation of the blue curves by a zero-mean Gaussian.

observable. One can see that by introducing an independence assumption between the segments, the resulted distribution resembles the distribution in the case of two source observations, illustrated by the red dashed ellipses in Fig.B.2 b, Fig.B.2 a, respectively.



(a). Two sources are observable.



(b). A single source is observable.

Fig. B.2 In this figure the measurements samples of two segments with radius $l = 1$ are illustrated by the blue points in two scenarios: (a) both ends of the segments are observable and (b) a single end is observable. Also, the figure compares the measurement likelihood obtained by the general model (4.98) and the independent model (4.99) illustrated by the black and red dashed ellipses respectively.