

EXTENDED TARGET TRACKING OF CONVEX  
POLYTOPE SHAPES

EXTENDED TARGET TRACKING OF CONVEX  
POLYTOPE SHAPES WITH MANEUVERS AND CLUTTER

By

PRABHANJAN MANNARI,

B.Tech (Electronics and Communication Engineering)

A Thesis Submitted to the School of Graduate Studies  
in the Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in

Electrical and Computer Engineering

McMaster University

Hamilton, Ontario

Doctor of Philosophy (2023)  
Electrical and Computer Engineering  
McMaster University  
Hamilton, Ontario, Canada

TITLE: Extended Target Tracking of Convex Polytope Shapes with Maneuvers  
and Clutter

AUTHOR:

Prabhanjan Mannari,  
B.Tech (Electronics and Communication Engineering)

SUPERVISOR:

Dr. Thiagalingam Kirubarajan  
Professor, ECE,  
McMaster University, ON, Canada

SUPERVISOR:

Dr. Ratnasingham Tharmarasa  
Professor, ECE,  
McMaster University, ON, Canada

SUPERVISORY COMMITTEE MEMBERS:

Dr. Sorina Dumitrescu  
Professor, ECE,  
McMaster University, ON, Canada

Dr. Timothy R. Field  
Professor, ECE,  
McMaster University, ON, Canada

NUMBER OF PAGES: xxi, 180

# Abstract

High resolution sensors such as automotive radar and LiDAR have become prevalent in target tracking applications in recent times. Data from such sensors demands extended target tracking in which, the shape of the target is to be estimated along with the kinematics. Several applications benefit from extended target tracking, for example, autonomous vehicles and robotics.

This thesis proposes a different approach to extended target tracking compared to existing literature. Instead of a single shape descriptor to describe the entire target shape, different parts of the extended target are assumed to be distinct targets constrained by the target rigid body shape. This formulation is able to handle issues such as self-occlusion and clutter which, are not addressed sufficiently in literature.

Firstly, a framework for extended target tracking is developed based on the formulation proposed. Using 2D convex hull as a shape descriptor, an algorithm to track 2D convex polytope shaped targets is developed. Further, the point target Probabilistic Multiple Hypotheses Tracker (PMHT) is modified to derive an extended target PMHT (ET-PMHT) equations to track 3D convex polytope shapes, using a Delaunay triangulation to describe the shape. Finally, the approach is extended to handle target maneuvers, as well as, clutter and measurements from the interior of the target.

In all three cases, the issue of self-occlusion is considered and the algorithms are

still able to effectively capture the target shape. Since the true target center may not be observable, the shape descriptor abandons the use of target center in the state, and the shape is described by its boundary alone. The shape descriptors also support addition and deletion of faces, which is useful for handling newly visible parts of the target and clutter, respectively.

The algorithms proposed have been compared with the existing literature for various scenarios, and it is seen that the proposed algorithms outperform, especially in the presence of self-occlusion.

# Acknowledgements

I would like to use this opportunity to thank everyone who helped me in this endeavour. Firstly, I am grateful to my supervisors Prof. T. Kirubarajan and Dr R. Tharmarasa for their continued support through the years, both financially and in technical matters. I also found their courses extremely helpful through the course of my research.

I appreciate the timely reviews and suggestions from the committee members Prof. S. Dumitrescu and Prof. T. Field. I would also like to thank my lab mates who have made this journey memorable. I am pleased to be a part of the Electrical and Computer Engineering department and, I thank the faculty for offering insightful courses and the administrative staff, especially Ms. C. Gies, who helped me navigate through the department.

Finally, I like to express my sincere gratitude and thanks to my mother, father and my sister for their consistent assurance and support throughout these years.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>v</b>
<b>List of Abbreviations</b>	<b>xi</b>
<b>List of Figures</b>	<b>xiii</b>
<b>List of Tables</b>	<b>xx</b>
<b>1 Introduction</b>	<b>1</b>
1.1 A brief review of Extended Target Tracking . . . . .	1
1.2 Literature Review . . . . .	3
1.3 Theme and Objectives of the Dissertation . . . . .	6
1.4 Summary of Enclosed Articles . . . . .	8
1.4.1 Paper I (Chapter 2) . . . . .	8
1.4.2 Paper II (Chapter 3) . . . . .	9
1.4.3 Paper III (Chapter 4) . . . . .	10
Bibliography . . . . .	11
<b>2 Extended Target Tracking under Multitarget Tracking Frame- work for Convex Polytope shapes</b>	<b>14</b>

2.1	Introduction . . . . .	16
2.2	Problem Description . . . . .	22
2.2.1	Target model . . . . .	23
2.2.2	Measurement Model . . . . .	25
2.3	Probability Update . . . . .	28
2.4	Extended Target Tracking in Multitarget Tracking Framework . . .	33
2.5	Implementation . . . . .	36
2.5.1	Determining the measurement function - Association of mea- surement to source location . . . . .	37
2.5.2	Convex hull and open convex hull . . . . .	41
2.5.3	Initialization of velocity and shape . . . . .	42
2.5.4	Prediction . . . . .	45
2.5.5	Update . . . . .	46
2.6	Results . . . . .	50
2.6.1	Gaussian Process (GP) . . . . .	52
2.6.2	Scenario . . . . .	53
2.6.3	Different levels of measurement noise . . . . .	55
2.6.4	Different levels of average number of measurements . . . . .	60
2.6.5	Performance for different shapes . . . . .	63
2.6.6	Discussion and Remarks . . . . .	66
2.7	Conclusions and Future Work . . . . .	67
	Bibliography . . . . .	69

**3 Extended Target Tracking using ET-PMHT for 3D Convex Poly-  
tope Shapes with Partial Visibility 73**

3.1	Introduction . . . . .	75
3.2	Problem Description . . . . .	80
3.2.1	Target Model . . . . .	80
3.2.2	Measurement Model . . . . .	83
3.3	MTT Framework for single extended target tracking . . . . .	86
3.4	Solution . . . . .	87
3.4.1	Determination of the measurement function . . . . .	87
3.4.2	Expectation Maximization for static shape estimation: ET- PMHT initialization . . . . .	89
3.4.3	Expectation Maximization for dynamic state estimation: ET- PMHT update . . . . .	92
3.4.4	Simplification of the convex hull . . . . .	99
3.5	Algorithm . . . . .	100
3.5.1	Initialization . . . . .	100
3.5.2	Update . . . . .	101
3.6	Results . . . . .	104
3.6.1	3D Gaussian Process . . . . .	105
3.6.2	Scenario . . . . .	107
3.6.3	Different levels of measurement noise when all parts of the target are visible . . . . .	109
3.6.4	Different levels of measurement noise with partial visibility .	111
3.6.5	Different levels of average number of measurements per unit area with partial visibility . . . . .	114
3.7	Conclusions and Future work . . . . .	116
3.8	Appendix - Derivatives of the auxiliary function . . . . .	117

Bibliography . . . . . 118

**4 ET-PMHT for Tracking 3D Extended Targets with Maneuvers  
and Clutter . . . . . 121**

4.1 Introduction . . . . . 123

4.2 Problem Description . . . . . 128

    4.2.1 Target Model . . . . . 129

    4.2.2 Measurement Model . . . . . 132

4.3 Solution . . . . . 134

    4.3.1 Extended Target Tracking Framework . . . . . 134

    4.3.2 Determination of the measurement function . . . . . 137

    4.3.3 Expectation Maximization . . . . . 138

4.4 Derivation of the ET-PMHT equations . . . . . 139

    4.4.1 ET-PMHT for initialization with clutter . . . . . 139

    4.4.2 State extension . . . . . 140

    4.4.3 ET-PMHT for update with clutter . . . . . 141

    4.4.4 ET-PMHT with fixed shape . . . . . 146

    4.4.5 ET-PMHT shape update . . . . . 148

    4.4.6 Face to Face association . . . . . 149

4.5 Algorithm . . . . . 150

    4.5.1 Initialization . . . . . 150

    4.5.2 Update . . . . . 154

4.6 Results . . . . . 155

    4.6.1 3D Gaussian Process . . . . . 157

    4.6.2 Scenario . . . . . 158

4.6.3	Different levels of measurement noise covariance . . . . .	160
4.6.4	Different levels of average number of measurements per unit area . . . . .	163
4.6.5	Different levels of clutter . . . . .	166
4.7	Conclusions and future work . . . . .	169
4.8	Appendix - Derivation of the Jacobian of the transition function . .	170
	Bibliography . . . . .	171
<b>5</b>	<b>Conclusions and Future Work</b>	<b>175</b>
5.1	Conclusions . . . . .	175
5.1.1	Challenges and Approaches . . . . .	176
5.2	Future Work . . . . .	179

## Abbreviations

**CH** Convex Hull

**CT** Co-ordinated Turn

**DT** Delaunay Triangulation

**EKF** Extended Kalman Filter

**EM** Expectation Maximization

**ET** Extended Target

**ET-PMHT** Extended Target - Probabilistic Multiple Hypotheses Tracker

**FM** Face Management

**FOV** Field of View

**GP** Gaussian Process

**IoU** Intersection over Union

**JPDA** Joint Probabilistic Data Association

**KF** Kalman Filter

**LiDAR** Light Detection and Ranging

**LRKF** Linear Regression Kalman Filter

**MHT** Multiple Hypotheses Tracker

**MTT** Multitarget Tracking

**NCV** Nearly Constant Velocity

**PDA** Probabilistic Data Association

**PMHT** Probabilistic Multiple Hypotheses Tracker

**PSD** Positive Semi Definite

**RHM** Random Hypersurface Model

**RM** Random Matrix

**RMSE** Root Mean Square Error

**UKF** Unscented Kalman Filter

# List of Figures

1.1	Example of measurements from an extended target [1]. The measurements are shown in red and the sensor location is shown in blue	2
2.1	The effect of sensor target geometry on number of measurements obtained	26
2.2	Example of measurement from a target with the sensor at $[0\ 0]$ . The visible faces are green and the measurements are red in color	28
2.3	Visual representation of Equation (2.13)	31
2.4	Flowchart of the frameworks	37
2.5	Target shape and sample trajectory	54
2.6	RMSE center for $\sigma_x = \sigma_y = 0.1m$ and $\rho = 100m^{-1}$ when all faces are visible.	55
2.7	RMSE velocity for $\sigma_x = \sigma_y = 0.1m$ and $\rho = 100m^{-1}$ when all faces are visible.	56
2.8	IoU results for $\sigma_x = \sigma_y = 0.1m$ and $\rho = 100m^{-1}$ when all faces are visible.	56

2.9	Estimates of the target shape at different times with $\sigma_x = \sigma_y = 0.1m$ and $\rho = 100m^{-1}$ . The true target shape is given in blue, and the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green . . . . .	56
2.10	RMSE center for different levels of measurement noise. Results for the proposed approach are on the left and the GP results are on the right . . . . .	57
2.11	RMSE velocity for different levels of measurement noise. Results for the proposed approach are on the left and the GP results are on the right . . . . .	57
2.12	IoU results for different levels of measurement noise. Results for the proposed approach are on the left and the GP results are on the right	58
2.13	Estimates of the target shape at different times for $\sigma_x = \sigma_y = 0.1m$ , $\rho = 100m^{-1}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green. . . . .	58
2.14	RMSE center for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right	60
2.15	RMSE velocity for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right	60
2.16	IoU results for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right	61

2.17	Estimates of the target shape at different times $\sigma_x = \sigma_y = 0.1m$ , $\rho = 40m^{-1}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green . . . . .	61
2.18	Rectangle, Pentagon, Hexagon, and custom target shapes . . . . .	63
2.19	RMSE center for different shapes. Results for the proposed approach are on the left and the GP results are on the right . . . . .	63
2.20	RMSE velocity for different shapes. Results for the proposed approach are on the left and the GP results are on the right . . . . .	64
2.21	IoU results for different shapes. Results for the proposed approach are on the left and the GP results are on the right . . . . .	64
2.22	Estimates of the target shape at different times $\sigma_x = \sigma_y = 0.1m$ , $\rho = 100m^{-1}$ for rectangle target shape. The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green . . . . .	65
2.23	Estimates of the target shape at different times $\sigma_x = \sigma_y = 0.1m$ , $\rho = 100m^{-1}$ for pentagon target shape. The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green . . . . .	65

2.24	Estimates of the target shape at different times for $\sigma_x = \sigma_y = 0.1m$ , $\rho = 100m^{-1}$ and hexagon target shape. The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green . . . . .	66
3.1	Example of self-occlusion. Visible faces are transparent and faces that are not visible are shaded. [1] . . . . .	76
3.2	Example of a target as a cube and its Delaunay triangulation . . . . .	81
3.3	Example of measurements from visible faces of the target shape . . . . .	85
3.4	Flowchart of the frameworks . . . . .	87
3.5	Target shape . . . . .	107
3.6	RMSE center for different levels of measurement noise when all faces are visible. Results for the proposed approach are on the left and the GP results are on the right . . . . .	110
3.7	RMSE velocity for different levels of measurement noise when all faces are visible. Results for the proposed approach are on the left and the GP results are on the right . . . . .	110
3.8	IoU results for different levels of measurement noise when all faces are visible. Results for the proposed approach are on the left and the GP results are on the right . . . . .	111

3.9	Estimates of the target shape at different times when all faces are visible for $\sigma_x = \sigma_y = \sigma_z = 0.1m$ , $\rho_N = 0.5m^{-2}$ . The true target shape is given in cyan, the measurements at the current time are given in red. The estimate using the proposed algorithm is in blue and the GP estimate is in green. . . . .	111
3.10	RMSE center for different levels of measurement noise with partial visibility. Results for the proposed approach are on the left and the GP results are on the right . . . . .	113
3.11	RMSE velocity for different levels of measurement noise with partial visibility. Results for the proposed approach are on the left and the GP results are on the right . . . . .	113
3.12	IoU results for different levels of measurement noise with partial visibility. Results for the proposed approach are on the left and the GP results are on the right . . . . .	113
3.13	Estimates of the target shape at different times $\sigma_x = \sigma_y = \sigma_z = 0.1m$ , $\rho_N = 0.1m^{-2}$ . The true target shape is given in cyan, the measurements at the current time are given in red. The estimate using proposed algorithm is in blue and the GP estimate is in green	114
3.14	RMSE center for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right	115
3.15	RMSE velocity for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right	115
3.16	IoU results for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right	115

4.1	Example of measurements from the target along with clutter. . . . .	135
4.2	Target Shape . . . . .	159
4.3	Sample Trajectory . . . . .	160
4.4	RMSE center for different levels of measurement noise covariance. Results for the proposed approach are on the left and the GP results are on the right . . . . .	161
4.5	RMSE velocity for different levels of measurement noise covariance. Results for the proposed approach are on the left and the GP results are on the right . . . . .	161
4.6	IoU results for different levels of measurement noise covariance. Re- sults for the proposed approach are on the left and the GP results are on the right . . . . .	162
4.7	Estimates of the target shape at different times $\sigma_x = \sigma_y = \sigma_z =$ $0.1m$ , $\rho = 3m^{-2}$ and $\rho_C = 0.1m^{-3}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using proposed algorithm is in black. . . . .	162
4.8	RMSE center for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right	164
4.9	RMSE velocity for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right	164
4.10	IoU results for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right	165

4.11	Estimates of the target shape at different times $\sigma_x = \sigma_y = \sigma_z = 0.1m$ , $\rho = 2m^{-2}$ and $\rho_C = 0.1m^{-3}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using proposed algorithm is in black. . . . .	165
4.12	RMSE center for different levels of clutter. Results for the proposed approach are on the left and the GP results are on the right . . . .	167
4.13	RMSE velocity for different levels of clutter. Results for the proposed approach are on the left and the GP results are on the right .	167
4.14	IoU results for different levels of clutter. Results for the proposed approach are on the left and the GP results are on the right . . . .	167
4.15	Estimates of the target shape at different times $\sigma_x = \sigma_y = \sigma_z = 0.1m$ , $\rho = 3m^{-2}$ and $\rho_C = 0.15m^{-3}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using proposed algorithm is in black. . . . .	168

# List of Tables

2.1	Table of runtimes for different average levels of measurements . . .	62
-----	--	----

# Declaration of Academic Achievement

I, Prabhanjan Mannari, declare that this thesis titled, **Extended Target Tracking of Convex Polytope Shapes with Maneuvers and Clutter**, and works presented in it are my own. I confirm that analytical and computational work has been carried out solely by Prabhanjan Mannari, with the advice and guidance provided by the academic supervisors Prof. T. Kirubarajan and Prof. R. Tharmarasa. Information that is presented from outside sources, which has been used for analysis and discussion, has been cited when appropriate, and all the other materials are solely the work of the author.

# Chapter 1

## Introduction

### 1.1 A brief review of Extended Target Tracking

Traditional target tracking algorithms assume point targets and a vast amount of literature exists for point multitarget tracking, which includes filtering, data association, validation, track initialization and track management. Point targets are those whose extent occupies at most a single resolution cell of the sensor, and hence each target can generate at most a single measurement per frame. Recently, high resolution sensors such as automotive radar and LiDAR (Light Detection and Ranging) have become more prevalent in target tracking applications. The targets of interest for such sensors occupy multiple sensor resolution cells and can generate multiple measurements in a single frame. Such targets are termed as extended targets and have a finite extent compared to the sensor resolution cell as opposed to point targets that have zero extent. An example of measurements generated by an extended target is given in Figure 1.1. The measurements from an extended target contain information about the shape of the target, and hence the shape as

well as the kinematics of the target are to be estimated using measurements over time. The size of the resolution cells of the sensor can vary with the distance from the sensor. The same target can be considered as a point target or an extended target depending on the sensor-target geometry.

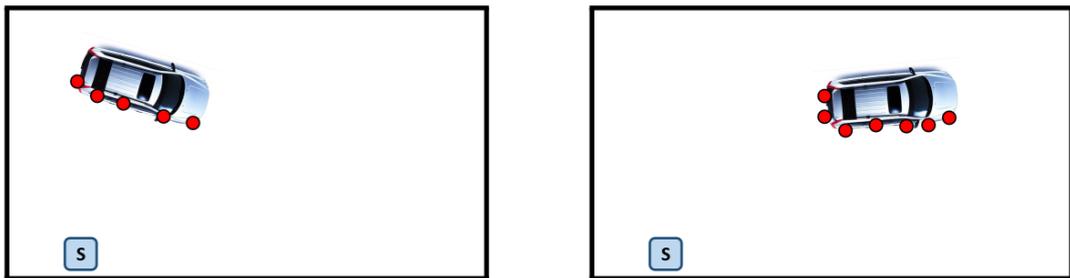


FIGURE 1.1: Example of measurements from an extended target [1]. The measurements are shown in red and the sensor location is shown in blue

Extended target tracking provides a shape estimate of the target and has several downstream applications, for example, classification of targets, obstacle detection and path planning to name a few. Autonomous vehicle and robotics applications use high resolution sensors for close range scenarios and particularly benefit from extended target tracking.

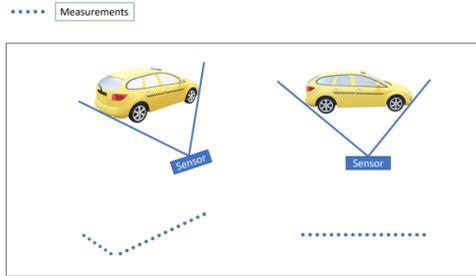
One of the major challenges in extended target tracking is the joint uncertainty in the shape and the kinematics of the target, i.e. the same set of measurements can be generated over time by targets with distinct combinations of shapes and kinematics. As such, the target state needs to include a shape descriptor to describe the target shape. Filtering for estimation of the target shape using the measurements needs to be developed depending on the shape descriptor used.

The source location for the measurement on the target shape is unknown as well and is termed as measurement-origin uncertainty.

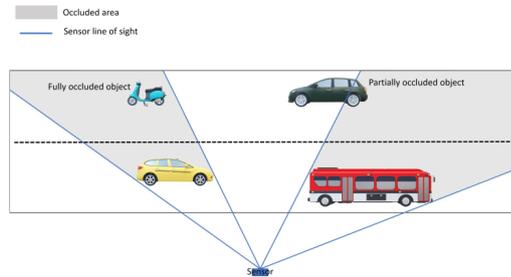
Occlusion is yet another challenge in extended target tracking. Measurements are generated only from the parts of the target visible to the sensor. Parts or faces of the target to which the sensor has a direct line of sight are visible to the sensor. Even in the case of a single target, the entire target shape is not visible in the same frame, particularly in the scenario with a single sensor. The case when the line of sight to some parts of the target is blocked by the target itself is termed as self-occlusion, and an example is given in Figure 1.2a. The problem is further exacerbated in case of multiple targets as seen in Figure 1.2b. This case, when the line of sight of the sensor to an extended target is occluded/blocked by another extended target is termed as mutual occlusion. Currently, in the thesis, the scenario with a single extended target and a single sensor is considered with self-occlusion. The objective of an extended target tracking algorithm is to estimate the entire shape of the target even when the measurements occur from different parts of the target in different frames due to self-occlusion.

## **1.2 Literature Review**

Several approaches for extended target tracking have been proposed in literature, and they can be categorized based on the shape descriptor used. Simple shapes such as ellipses and rectangles have been used in [3],[4] and the parameters such as length/breadth or major/minor axes are to be estimated. The Gaussian Process (GP) approach used in [5] and [6] uses a star-convex shape descriptor that can



(A) Example of self-occlusion [2].



(B) Example of mutual occlusion [2]

describe a variety of complicated shapes. The Random Hypersurface Model (RHM) used in [7] supports different shape descriptors including ellipse, rectangle and polygon to describe more complex shapes. A shape model using splines has been presented in [8] and [9].

The Random Matrix (RM) approach first proposed in [10] uses a symmetric positive semi definite (PSD) matrix  $X_S$  to describe the shape of the target with  $x_c$  as the center. More specifically, the measurements from the target are obtained around the target center  $x_c$  with covariance  $X_S$ . The shape model is appropriate for ellipsoidal shapes. For example, even a rectangular shaped target can generate measurements closer to an ellipsoid if the measurement noise covariance is large enough compared to the shape. The shape matrix  $X_S$  is assumed to be Inverse

Wishart distributed since it is a conjugate prior and an analytical Bayesian solution is derived. The Random Matrix method is extended in [11] with multiple ellipsoids to represent non-ellipsoidal shapes more effectively. Multipath detections are considered in a terrain constrained environment in [12] using the Random Matrix model. Further extensions to the Random Matrix model can be found in [1]. The Random Matrix can estimate the general features of a complicated, however it is unable to capture finer features of the shape.

The Gaussian Process proposed in [5] for 2D and [6] for 3D uses a radial basis function to describe star-convex shapes. The radii follow a Gaussian Process with a kernel function  $\kappa$  to describe the covariance between the radii for different angles. An Extended Kalman Filter (EKF) is used to handle the nonlinearity in the joint filtering of the shape and kinematics. The Gaussian Process approach has been extended in [13] to handle the scenario when a target changed from a point target to an extended target (or vice-versa) due to the varying sensor-target geometry by using a Poisson rate for the number of measurements from the target. The article [14] includes clutter by using a Probabilistic Data Association (PDA) along with the Gaussian Process shape descriptor.

The Random Hypersurface Model approach [7] uses a pseudo-measurement equation  $\phi(X_S, Z) = 0$ , where  $X_S$  is the shape descriptor,  $Z$  are the measurements and  $\phi$  is a distance metric between the measurements and the source location on the target shape. The objective is to estimate the shape that minimizes the distance metric.

Various models to describe a variety of shapes have been proposed in literature,

however the problem of self-occlusion has not been addressed sufficiently. This is a major limitation and visibility issues need to be included for more realistic assumptions. The models used in literature generally include the target center as a part of the state, whereas the true target center may be unobservable, for example, when the target is only partially visible across frames. The existing methods use a fixed number of parameters to describe the target shape. It is useful to vary the number of parameters used to describe the shape as new parts of the target become visible over time.

This thesis aims to relax these assumptions, particularly that of self-occlusion, to develop algorithms suitable for more realistic scenarios. To this end, a framework for extended target tracking is proposed and algorithms to track 2D and 3D convex polytope shapes are developed.

### **1.3 Theme and Objectives of the Dissertation**

In compliance with the terms and regulations of McMaster University, this dissertation has been written in the sandwich thesis format by assembling three articles. These articles represent the independent research performed by the author of this dissertation, Prabhanjan Mannari.

The focus of the thesis is to develop a general framework for extended target tracking and implement algorithms for tracking convex polytope shaped targets with self-occlusion in the presence of maneuvers and clutter. The main highlights of the contribution of the thesis are as follows :

1. A framework for extended target tracking based on the point multitarget tracking framework was proposed (Paper I) and extended to 3D (Paper II and Paper III). Under the framework, different parts of a single extended target are treated as distinct targets constrained by the target rigid body shape. This allows for handling of self-occlusion and clutter by associating measurements to appropriate parts of the target. The framework effectively transforms the extended target tracking problem to a complex multitarget tracking problem with closely spaced targets (the faces of the target may, in fact, have common edges as well).
2. A shape descriptor for convex polytope shapes has been proposed for 2D (Paper I) and 3D (Paper II and Paper III). The model uses a convex hull representation to describe the shape only by its boundary, and the center is not included in the shape model. The triangular mesh representation used for 3D convex polytope shapes adopted in the thesis is widely used in engineering [15].
3. The self-occlusion issue is handled for 2D case (Paper I) as well as for the 3D case (Paper II and Paper III) by associating the measurements only to the visible parts of the target.
4. The point target Probabilistic Multiple Hypotheses Tracker (PMHT) has been modified to develop an extended target PMHT (ET-PMHT) for the shape triangulations proposed (Paper II). The ET-PMHT has further been extended to include clutter and to handle target maneuvers (Paper III).
5. Face management is proposed under the extended target tracking framework

that can be used to delete faces using a quality parameter and, proves to be useful especially in the presence of clutter (Paper I, Paper II and Paper III).

6. The framework supports adding new faces to the shape estimate as measurements are obtained from parts of the target that were previously not visible or as new features of the target become observable (Paper I, Paper II and Paper III).

## **1.4 Summary of Enclosed Articles**

A summary of the papers included in the thesis is given below:

### **1.4.1 Paper I (Chapter 2)**

Prabhanjan Mannari, Ratnasingham Tharmarasa and Thiagalingam Kirubarajan "Extended Target Tracking under Multitarget Tracking Framework for Convex Polytope shapes", Accepted by Elsevier Signal Processing journal as of November 6 2023.

Preface: The problem of tracking a 2D convex polytope shaped target is addressed in the paper. The target is assumed to be a rigid body with known dynamics, such as nearly constant velocity (NCV) model. A convex hull shape model is developed using a directed set of vertices to describe the shape. An extended target tracking framework is proposed based on the points multitarget tracking framework, in which different parts/faces of a single extended are assumed to be distinct targets.

Under this framework, the visible faces of the estimate can be identified and the measurements can be associated only to those faces, thus handling the problem of self-occlusion. The source location for a measurement is estimated as the closest point on the shape estimate to the measurement and nearest neighbour is used for associating measurements to faces of the estimate. The updated shape and kinematics are estimated using a Kalman Filter (KF) with uncertainty in the source locations, i.e. the true source locations are unknown and an estimate is used instead. New faces are initialized using the measurements falling outside the validation region of the estimate. This is necessary as new measurements may occur from parts of the target that were previously not visible due to self-occlusion. Face management step is proposed to delete erroneous faces. The performance of the proposed algorithm is compared with 2D Gaussian Process for various scenarios using simulations. The proposed algorithm is able to effectively estimate the entire target shape even in the presence of self-occlusion.

### **1.4.2 Paper II (Chapter 3)**

Prabhanjan Mannari, Ratnasingham Tharmarasa and Thiagalingam Kirubarajan "3D Extended Target Tracking using ET-PMHT for Convex Polytope Shapes with Partial Visibility" Ready to be submitted to a journal as of December 8, 2023.

Preface: This article discusses the problem of tracking a single 3D extended target, even when the target is only partially visible. The target is assumed to have a convex polytope shape and known dynamics, such as nearly constant velocity (NCV) model. The target shape is represented using its vertices and a Delaunay triangulation that divides the surface of the target shape into non-overlapping triangles.

Based on the extended target tracking framework proposed in our previous work, each of the face triangulations is assumed to be separate targets constrained by the target shape. The point target Probabilistic Multiple Hypotheses Tracker (PMHT) is modified to develop an extended target PMHT (ET-PMHT). The equations for ET-PMHT joint association and filtering are derived for a single frame update for the general case and for the case where the shape is unchanged across the update. The mode with constant shape estimate proves to be useful to maintain a good shape estimate that may otherwise degrade due to further measurements. Face initialization and face management techniques are used to add new faces and delete erroneous faces, respectively. The proposed algorithm is compared with the 3D Gaussian Process for different scenarios, and it is found that the proposed algorithm is able to track the target, even when it is only partially visible.

### **1.4.3 Paper III (Chapter 4)**

Prabhanjan Mannari, Ratnasingham Tharmarasa and Thiagalingam Kirubaran "ET-PMHT for Tracking 3D Extended Targets with Maneuvers and Clutter". Ready to be submitted to a journal as of December 8, 2023.

Preface: The problem of tracking a maneuvering extended target in 3D is discussed in this article. The visibility issue of self-occlusion and the effect of clutter as well as measurements from the interior of the target are included in the scenario. The target has a convex polytope shape and maneuvering dynamics such as coordinated turn with an unknown turn rate. Extended target PMHT (ET-PMHT) equations are derived under the framework developed in our previous works with

linearization to handle the nonlinear transition dynamics. The self-occlusion problem is handled by associating the current measurements only to the visible faces of the estimate. The effect of clutter is included in the ET-PMHT equations and face management is used to delete erroneous due to clutter. The algorithm supports adding new faces as parts of the target that were previously not visible, become visible over time. The performance of the proposed method is compared with the 3D Gaussian process for various scenarios and RMSE of the center, RMSE of the velocity and IoU metrics are used to quantify the performance.

## Bibliography

- [1] Karl Granstrom, Marcus Baum, and Stephan Reuter. Extended object tracking: Introduction, overview and applications, 2017.
- [2] Mehrnoosh Heidarpour. Extended target tracking methods in modern sensor applications. 2020. doi: <http://hdl.handle.net/11375/25713>.
- [3] Marcus Baum and Uwe Hanebeck. Tracking an extended object modeled as an axis-aligned rectangle. pages 2422–2427, 01 2009.
- [4] Shishan Yang and Marcus Baum. Tracking the orientation and axes lengths of an elliptical extended object. *IEEE Transactions on Signal Processing*, 67(18):4720–4729, 2019. doi: 10.1109/TSP.2019.2929462.
- [5] Niklas Wahlstrom and Emre Ozkan. Extended target tracking using gaussian processes. *IEEE Transactions on Signal Processing*, 63(16):4165–4178, 2015. doi: 10.1109/TSP.2015.2424194.

## BIBLIOGRAPHY

---

- [6] Murat Kumru and Emre Ozkan. Three-dimensional extended object tracking and shape learning using gaussian processes. *IEEE Transactions on Aerospace and Electronic Systems*, 57(5):2795–2814, oct 2021. doi: 10.1109/taes.2021.3067668. URL <https://doi.org/10.1109/taes.2021.3067668>.
- [7] Antonio Zea, Florian Faion, Marcus Baum, and Uwe D. Hanebeck. Level-set random hypersurface models for tracking nonconvex extended objects. *IEEE Transactions on Aerospace and Electronic Systems*, 52(6):2990–3007, 2016. doi: 10.1109/TAES.2016.130704.
- [8] Abdullahi Daniyan, Sangarapillai Lambotharan, Anastasios Deligiannis, Yu Gong, and Wen-Hua Chen. Bayesian multiple extended target tracking using labeled random finite sets and splines. *IEEE Transactions on Signal Processing*, 66(22):6076–6091, 2018.
- [9] Hauke Kaulbersch, Jens Honer, and Marcus Baum. A cartesian b-spline vehicle model for extended object tracking. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 1–5, 2018. doi: 10.23919/ICIF.2018.8455717.
- [10] Johann Wolfgang Koch. Bayesian approach to extended object and cluster tracking using random matrices. *IEEE Transactions on Aerospace and Electronic Systems*, 44(3):1042–1059, 2008. doi: 10.1109/TAES.2008.4655362.
- [11] Jian Lan and X. Rong Li. Tracking of maneuvering non-ellipsoidal extended object or target group using random matrix. *IEEE Transactions on Signal Processing*, 62(9):2450–2463, 2014. doi: 10.1109/TSP.2014.2309561.

## BIBLIOGRAPHY

---

- [12] Ben Liu, Ratnasingham Tharmarasa, Rahim Jassemi, Daly Brown, and Thia Kirubarajan. Extended target tracking with multipath detections, terrain-constrained motion model and clutter. *IEEE Transactions on Intelligent Transportation Systems*, 22(11):7056–7072, 2021. doi: 10.1109/TITS.2020.3001174.
- [13] Xu Tang, Mingyan Li, Ratnasingham Tharmarasa, and Thiagalingam Kirubarajan. Seamless tracking of apparent point and extended targets using gaussian process pmht. *IEEE Transactions on Signal Processing*, 67(18):4825–4838, 2019. doi: 10.1109/TSP.2019.2932873.
- [14] Yunfei Guo, Yong Li, Ratnasingham Tharmarasa, Thiagalingam Kirubarajan, Murat Efe, and Bahadir Sarikaya. Gp - pda filter for extended target tracking with measurement origin uncertainty. *IEEE Transactions on Aerospace and Electronic Systems*, 55(4):1725–1742, 2019. doi: 10.1109/TAES.2018.2875555.
- [15] S. S. Rao. *The finite element method in engineering*. Elsevier/Butterworth Heinemann, Amsterdam ;, 4th ed. edition, 2005. ISBN 1-280-96441-3.

## Chapter 2

# Extended Target Tracking under Multitarget Tracking Framework for Convex Polytope shapes

The content of this chapter has been published in the Elsevier Signal Processing journal. The paper can be found here -

---

**Mannari, Prabhanjan** and Tharmarasa, Ratnasingham and Kirubarajan, Thiagalingam, "Extended Target Tracking Under Multitarget Tracking Framework for Convex Polytope Shapes", Signal Processing, Volume 217, 2024, <https://doi.org/10.1016/j.sigpro.2023.109321>.

---

# **Extended Target Tracking under Multitarget Tracking Framework for Convex Polytope shapes**

## **Abstract**

This paper discusses the problem of extended target tracking for a single 2D extended target with a convex polytopic shape and known dynamics. Extended targets are those that produce multiple measurements for a single frame. One of the major challenges in extended target tracking is the joint uncertainty in the shape and the kinematics of the target. Another challenge is the lack of visibility due to self-occlusion in targets with a finite extent (as opposed to zero extent for point targets). To address these challenges, we develop a framework for tracking single (or widely separated) extended targets. This framework is based on the existing point multitarget tracking framework by modeling different parts of an extended target as separate targets. An algorithm is developed using the proposed framework for tracking convex polytope-shaped targets. The proposed shape function consists only of the boundary of the target since the center may not be observable. The algorithm is capable of dynamically changing the number of parameters used to describe the shape as more parts of the target become visible over time. The performance of the algorithm is evaluated for various scenarios using root mean square error (RMSE) of velocity, center, and intersection over union (IoU) metrics. It is seen that the algorithm is able to handle the self-occlusion problem and estimate the whole target shape even when different parts of the target are visible at different frames, for various shapes, and for various conditions of measurement

noise covariance and number of measurements. New faces are added to the shape estimate as more parts of the target become visible. The algorithm is able to conserve the parts of the target that were visible in the previous frames but are no longer visible.

**Keywords:** *Extended Target, Convex hull, Data Association, Self-Occlusion*

## **2.1 Introduction**

Traditional target tracking algorithms assume that the targets have zero extent (i.e., they are point targets) and that a measurement from the target falls into at most one resolution cell of the sensor. Such algorithms are termed as 'point target tracking' algorithms. This is the scenario when the target size or extent is small enough to fit in a single resolution cell of the sensor. However, with the increasing availability of high-resolution sensors such as Light Detection And Ranging (LiDAR), these assumptions are no longer valid as measurements from a single target may occur from different resolution cells for the same frame. Hence, the targets need to be modeled as 'extended targets', and the shape (extent), as well as the kinematics of the targets, need to be estimated. Extended targets occupy multiple sensor resolution cells, and it is possible to obtain multiple measurements from an extended target for the same scan. The measurements from an extended target depend on the sensor-target geometry. Indeed, the same target can be considered as a point target or an extended target depending on the sensor-target geometry. The major challenge in tracking extended targets is estimating their shape while

handling the joint uncertainty between shape and kinematics, since the same set of measurements over time can be produced by targets with distinct shapes and kinematics.

Extended target tracking becomes relevant when high-resolution sensors are used such as automotive radar in transport applications and LiDAR in robotics applications. The shape and kinematics estimates of the objects in the environment can then be used for applications such as obstacle detection, drivable area detection, path planning and classification to name a few.

Several approaches have been proposed to handle the extended target tracking problem. In [1],[2],[3], the target shape is assumed to be a simple shape, such as a rectangle or an ellipse, whose parameters need to be estimated. The kinematic and shape parameters are decoupled and estimated separately. In [3], the elliptical shape is described by its center, its orientation, and the lengths of the semi-axes. A multiplicative error model has been used for the measurement equation. The Random Matrix (RM) model [1] extends the kinematic point target state with a positive semi-definite (PSD) matrix  $X$ , which represents the target shape as an ellipse. The kinematic state follows the Gaussian model, while the matrix that represents the extent follows the Inverse Wishart model. A closed-form expression for kinematic and shape update has been derived in both [1],[3]. In article [4], multiple random matrices are used to approximate and track a non-ellipsoidal shape more effectively than the standard RM approach. The RM model is extended in [5] to include multipath detections and clutter in a terrain-constrained environment. Probabilistic Data Association (PDA) is used to handle multipath detections and clutter, while a variational Bayesian technique is used to reduce the computational

complexity. Note that several other extensions of the RM model exist [6]. The Random Matrix approach uses a simple shape model for extended target tracking and is unable to effectively estimate complex shapes such as polygons. The shape estimate is always an ellipse and fails to capture the features of complicated target shapes.

Another approach to handle more complex shapes is to use a star convex model for the target shape. In [7], [8], a radial basis function is used to describe the target shape. The parameters (radii for different angles) are assumed to follow a Gaussian Process (GP) with a kernel function to determine the covariance between the parameters (radii) for different inputs (angles). The total state (including shape and kinematics) is estimated recursively using an Extended Kalman Filter (EKF). The complete formulation of the GP [7] is given in Section 2.6.1. The article [9] extends the standard GP model using a Poisson measurement rate for each basis point. The algorithm is able to handle clutter and missed detections using the Probabilistic Multiple Hypotheses Tracker (PMHT) for association. The tracker can simultaneously handle both extended and point targets. In [10], the GP model is extended to handle clutter using PDA, and a relation between point target PDA and the proposed extended target PDA is shown. The Gaussian Process approach is able to capture a large number of star convex shapes with the radial basis model, however the shape estimation depends on the choice of parameters for the kernel function used to describe the correlation between different parts of the target shape. The GP uses a single kernel function for the entire shape. Hence, if the target shape has both smooth parts (parts of the target with high correlation) and jagged parts (parts of the target with low correlation), the performance of the GP

is not optimal.

In [11], an approach to handle several different kinds of shapes is presented, termed as the Random Hypersurface Model (RHM). The measurements  $z$  are associated to source locations in the shape  $x^p$  using a shape distance function  $\phi$  to form a pseudo-measurement equation  $\phi(x^p, z) = 0$ . Simple shapes, such as a circle or an ellipse, can be used with this method. Complicated shapes can also be represented using a polygonal shape descriptor. Once the measurements are associated to their source locations, the state can be updated by using a Linear Regression Kalman Filter (LRKF) or an approximation of the likelihood to propagate the density using a particle filter. The GP and the RHM approaches can be used to incorporate measurements from the contour as well as the interior of the target using a scaling factor. Yet another model for the target shape is the spline model, described in [12] and [13]. [12] also addresses multiple extended targets using labeled random finite sets. A comprehensive review of the existing literature for extended target tracking is given in [6].

One of the major challenges in extended target tracking is self-occlusion. Depending on the sensor-target geometry, measurements may occur only from certain parts of the target that are visible to the sensor. As the target moves, the sensor-target geometry changes and hence, measurements can occur from different parts of the target at different times. This problem has not been sufficiently addressed in the existing literature. Most approaches to extended target tracking track the centroid of the target as a kinematic state. However, the true centroid of the target may be unobservable, depending on the measurements obtained. The target shape usually has a fixed number of parameters. It can be useful to vary the

number of parameters required to estimate the target as measurements occur from previously unseen parts of the target. In the existing literature, the entire set of measurements is used to update the target shape. However, the shape features of different parts of the target depend only on the respective local measurements.

In [14], a nonparametric maximum likelihood is introduced, where the likelihood is calculated entirely based on data. This motivates the choice of a nonparametric shape model, which is not limited by previously defined parameters. The article [15] proposes a method for maximum likelihood estimation for convex hull operation on the convex set to be estimated and noise under various assumptions. The choice of the convex polytope model is based on these references. A convex polytope is a generalization of a convex polygon (2D) or a convex polyhedron (3D) to higher dimensions. A 2D convex polytope (polygon) represented by its vertices and edges/faces is used in the current work. Consecutive vertices of the polytope form the edges/faces. The shape model is able to represent only the boundary of the target, abandoning the center. The convex polytope model can accommodate a variety of shapes while having a simple representation. The shape model is particularly relevant in applications where the target can be approximated by a convex polygon, for example - vehicles in automotive applications. A linear measurement model can be developed for the shape model. The shape model is non-parametric and is estimated entirely using the measurements. For example, an initial estimate with a single set of measurements can be the convex hull of the measurements.

A representation of partial target shape is required to handle the self-occlusion problem. Since the measurements occur from the visible parts of the target, an

association step is necessary to update only the visible parts using the local measurements they are associated to. A procedure to add additional faces is needed to refine or extend the target shape as more details or areas of the target become visible over time.

The issues listed above motivate using the point multitarget tracking framework and modifying it accordingly to handle the single extended target tracking problem. The individual faces of the convex polytope used to represent the extended target are treated as separate targets and hence the measurements can be associated only to the visible faces. Additional faces can be added for measurements falling outside the validation region, similar to Track Initialization. Erroneous faces can be deleted using a quality parameter similar to Track Management, which is a technique used in multitarget tracking to add new tracks, confirm existing tracks or delete existing tracks [16].

The major contributions of this paper are :

1. A generic framework is developed based on the existing point multitarget tracking framework to handle the single extended target tracking problem.
2. An algorithm is developed under the framework proposed to track 2D convex polytope-shaped targets. An iterative approach is used to handle the uncertainty in the source location of the measurements.
3. The center of the target is abandoned from the state since it may be unobservable. The shape is thus described by its boundary alone.

4. The algorithm is able to dynamically change the number of parameters used to describe the state by initializing new faces.
5. The measurements are associated with separate parts of the target and the associated sets of measurements are filtered separately, constrained by the relation between the adjacent faces. This ensures that the measurements contribute only to the local region of the target.
6. The self-occlusion problem is handled by determining the visible parts of the target and associating the measurements only to these parts.

The paper is organized as follows - Section 2.2 describes the problem with the target and measurement models. The theory for probability update is developed in Section 2.3. The extended target tracking framework for a single target is described in Section 2.4, and the algorithm implemented under the framework is described in Section 2.5. Results are presented in Section 2.6. Finally, Section 2.7 summarizes the conclusions derived from the study.

## **2.2 Problem Description**

The problem of tracking a single extended target (or widely separated targets) is discussed in the absence of clutter (usually outliers from clustering in a high-resolution scenario). An extended target has shape and kinematic features, which are to be estimated using measurements from the target over time. The target is assumed to be a rigid body with a convex polytopic shape. The target is assumed to have known dynamics, such as the nearly constant velocity (NCV) model. The

target state denoted by  $\mathbf{x}$  consists of the shape descriptor  $X_S$  and the target kinematics  $v$ ,

$$\mathbf{x} = \begin{bmatrix} X_S \\ v \end{bmatrix}. \quad (2.1)$$

### 2.2.1 Target model

The shape descriptor of the target  $X_S$  is represented (in 2D) by the vertices of the target shape (convex polytope) in the counterclockwise order. The actual target shape is obtained by taking the convex hull of the vertices. The faces of the target can be described by a pair of adjacent vertices (the last vertex is paired with the first vertex). Let the vertices of the shape be  $\left[ p_1^T \ p_2^T \ \dots \ p_{N_S}^T \right]^T$ , where  $N_S$  is the number of vertices of the target shape and each vertex  $p_i = \left[ p_i^x \ p_i^y \right]^T$ , where  $p_i^x$  and  $p_i^y$  are the  $x$  and  $y$  co-ordinates of the vertex. The target state is then given by -

$$X_S = \left[ p_1^x \ p_1^y \ p_2^x \ p_2^y \ \dots \ p_{N_S}^x \ p_{N_S}^y \right]^T, \quad v = \left[ v_x \ v_y \right]^T, \quad (2.2)$$

$$\mathbf{x} = \left[ p_1^x \ p_1^y \ p_2^x \ p_2^y \ \dots \ p_{N_S}^x \ p_{N_S}^y \ v_x \ v_y \right]^T. \quad (2.3)$$

The target dynamics, assuming the NCV model, and with  $k$  as the index for the time frames, is given by -

$$\mathbf{x}(k+1) = F(k)\mathbf{x}(k) + \Gamma(k)\nu(k), \quad \text{where } \nu(k) \sim \mathcal{N}(0, Q(k)). \quad (2.4)$$

Let  $T$  be the time between frames  $k$  and  $k+1$ . The transition matrix  $F(k)$  and the process noise covariance  $Q(k)$  are then given by -

$$F(k) = \begin{bmatrix} 1 & 0 & 0 & \dots & T & 0 \\ 0 & 1 & 0 & \dots & 0 & T \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0 & T \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad Q(k) = q \begin{bmatrix} \frac{1}{3}T^3 & \frac{1}{2}T^2 & 0 & 0 \\ \frac{1}{2}T^2 & T & 0 & 0 \\ 0 & 0 & \frac{1}{3}T^3 & \frac{1}{2}T^2 \\ 0 & 0 & \frac{1}{2}T^2 & T \end{bmatrix}, \quad (2.5)$$

$$\Gamma(k) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where  $q$  is the power spectral density of the process noise.

Note that this model is slightly different from the standard NCV model. The first  $2N_S$  elements of the state are the  $x$  and  $y$  co-ordinates denoting the position of the vertices and the last 2 elements are the velocities of the target in the  $x$  and  $y$  directions ( $v_x$  and  $v_y$  respectively). Since the target is a rigid body, all the vertices of the target have the same kinematics, i.e., the same shift for each time step. This is achieved by using a process noise  $\nu(k)$  of size  $4 \times 1$  and using matrix  $\Gamma(k)$ .

## 2.2.2 Measurement Model

Measurements are obtained from the faces of the target visible to the sensor. The measurements are assumed to occur only from the surface of the target. The visibility of a face of the target depends on the sensor-target geometry. For a convex polytope, a face is visible to the sensor if the center (mean) of the polytope and the sensor location are on opposite sides of the face.

The target is assumed to have a surface with a known average number of measurements per unit effective area -  $\rho$ . An effective area  $A_e$  is introduced to account for the variability in the number of measurements from a face due to sensor-target geometry. The effective area of a face  $A_e$  with respect to the sensor location is defined as :

$$A_e = A \times \sin\left(\frac{\alpha_S}{2}\right), \quad (2.6)$$

where  $A$  is the actual area of the face and  $\alpha_S$  is the angle subtended by the face at the sensor. The number of measurements from each visible face is assumed to be Poisson-distributed with the average number of measurements  $N_f$ ,

$$N_f = \rho \times A_e. \quad (2.7)$$

In reality, the expected number of measurements from a face is usually nonlinear outside a certain range and does not follow (2.7). The expected number of measurements for a face very close to the sensor may be zero instead. Such models can be included and handled in the Association and Face Management step of the framework.

The measurements are assumed to be uniformly distributed along the face. An example is shown in Figure (2.1), where the same face at different locations with respect to the sensor produces a different number of measurements. The face is of length  $10m$  and average number of measurements per unit effective area (length)  $\rho = 10m^{-1}$ . The half angle subtended at the face as well as the average number of measurements decreases when the face moves away from the sensor or is inclined with respect to the sensor's field of view (FOV), as in Figure (2.1c).

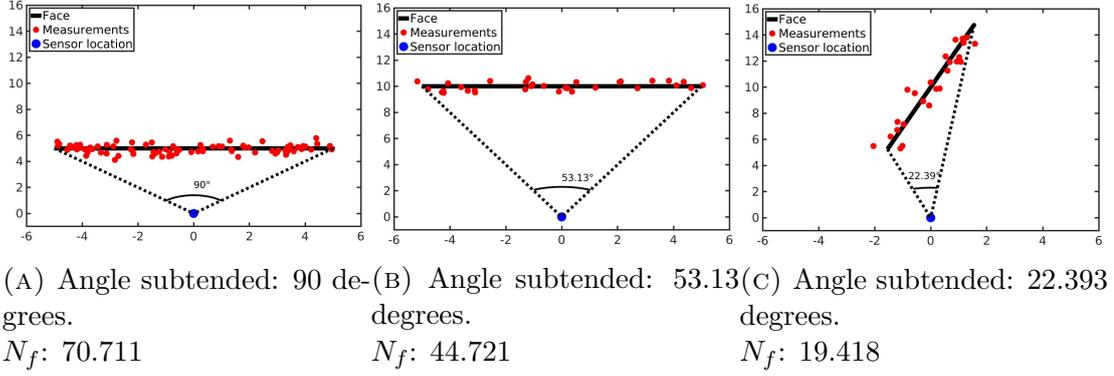


FIGURE 2.1: The effect of sensor target geometry on number of measurements obtained

A measurement  $z_i = \begin{bmatrix} z_i^x \\ z_i^y \end{bmatrix}$  occurs from an unknown source location  $s_i$  on the boundary of the target corrupted by Gaussian measurement noise with zero mean and known measurement noise covariance  $R$ .

$$z_i = s_i + w_i, \quad w_i \sim \mathcal{N}(0, R), \quad R = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}. \quad (2.8)$$

The source  $s_i$ , indexed by  $i$ , lies on the face  $(p_i, p_{i^*})$ , where  $(i^* = \text{mod}(i, N_S) + 1)$  and can be written as :

$$s_i = (1 - \lambda)p_i + \lambda p_{i^*}. \quad (2.9)$$

where  $0 \leq \lambda \leq 1$  is the parameter that determines the location of the source between the vertices. If  $N_i$  measurements occur from face  $(p_i, p_{i^*})$ , the total measurement set is written as :

$$z_{ij} = (1 - \lambda_{ij})p_i + \lambda_{ij}p_{i^*} + w_{ij} \quad \forall \quad i = 1 \text{ to } N_S \quad j = 1 \text{ to } N_i, \quad (2.10)$$

$$\text{where } w_{ij} \sim \mathcal{N}(0, R), \quad \lambda_{ij} \sim \mathcal{U}[0, 1], \quad N_i \sim \text{Poisson}(N_f^i). \quad (2.11)$$

$z_{ij}$  is the  $j$ th measurement from face  $(p_i, p_{i^*})$ .  $\lambda_{ij} \sim \mathcal{U}[0, 1]$  signifies that the measurements are uniformly distributed across the face and  $N_i \sim \text{Poisson}(N_f^i)$  implies that the number of measurements for the face is Poisson distributed with parameter  $N_f^i$  calculated according to Equation (2.7). It is assumed that the sensor is calibrated in advance using a technique such as [17] and all the measurement noise sources can be represented using zero mean Gaussian noise with measurement noise covariance  $R$ . An example of measurements obtained from the target is shown in Figure (2.2). Only two faces (shown in green) of the target are visible and measurements are obtained only from these faces. No measurements are obtained from the faces (shown in black) that are not visible. It can also be seen that the density of the measurements is lower for the visible face, which is longer since it is inclined and subtends a lower angle at the sensor.

In the current work, a simple model has been used for the distribution of the measurements along the face, with uniform distribution being used. In reality, the distribution of the measurements along the faces may be skewed, with measurements being concentrated towards one vertex of the face depending on the sensor

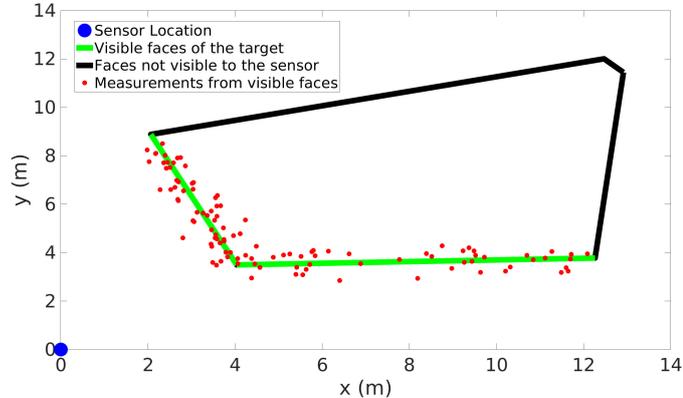


FIGURE 2.2: Example of measurement from a target with the sensor at  $[0\ 0]$ . The visible faces are green and the measurements are red in color

target geometry. When the inclination of the face with respect to the sensor is low, a uniform distribution is appropriate. However, faces with high inclination with respect to the sensor may have a distribution of measurements that is skewed along the face. More realistic models can be developed by assuming more complicated distributions for the probability of measurements being generated along the face. Using a Gaussian distribution with the mean at the mid-point of the face can be used to generate more measurements near the center of the face compared to the parts away from the center. To generate skewed distributions of measurements along the face, a Gaussian with mean away from the center can be used or a mixture of Gaussians can be used.

## 2.3 Probability Update

Given the measurements  $\mathcal{Z}^k = \{Z^1, Z^2, \dots, Z^k\}$  over time steps  $k = 1$  to  $K$  and a prior density  $p(\mathbf{x})$  of the state, the objective of a Bayesian solution is to estimate

the posterior of the state  $p(\mathbf{x}^k|\mathcal{Z}^k)$  for each  $k = 1, 2, \dots, K$ . The posterior is estimated recursively, i.e.,  $p(\mathbf{x}^K|\mathcal{Z}^K)$  is estimated using  $p(\mathbf{x}^{K-1}|\mathcal{Z}^{K-1})$  as the prior and measurements  $Z^K$  and  $p(\mathbf{x}^{K-1}|\mathcal{Z}^{K-1})$  is in turn estimated using  $p(\mathbf{x}^{K-2}|\mathcal{Z}^{K-2})$  as the prior and measurements  $Z^{K-1}$ , and so on. With Gaussian assumptions for the prior and posterior, the estimates of the probability densities for time step  $k$  can be described using the mean and covariance :  $\hat{\mathbf{x}}(k-1|k-1), P(k-1|k-1)$  for the prior and  $\hat{\mathbf{x}}(k|k), P(k|k)$  for the posterior, respectively. The state  $\mathbf{x}$  is described by a shape descriptor  $X_S$  with  $N_S$  vertices, and its kinematics.  $N_S$  can vary depending on the shape estimated.

At time  $k$ ,  $N_k$  measurements are obtained  $Z^k = \{z_1, z_2, \dots, z_{N_k}\}$  from faces described by the association variables  $\{a_1, a_2, \dots, a_{N_k}\}$ , where  $a_j \in \{1, 2, \dots, N_S\}$ . The association variable  $a_j = i$  denotes that the  $j^{\text{th}}$  measurement is associated to the  $i^{\text{th}}$  face. The exact measurement sources  $\{s_1, s_2, \dots, s_{N_k}\}$  and the association variables  $a_j$  are not observed. These sources can be described in terms of the state using a measurement source matrix  $\mathcal{H}$ , which is to be determined. The structure of  $\mathcal{H}$  is described below -

- (i) The number of rows in  $\mathcal{H}$  is compatible with the number of measurements, and the number of columns is compatible with the state shape and kinematics. Hence, the uncertainty about the number of vertices in the state shape is incorporated in  $\mathcal{H}$ .
- (ii)  $\mathcal{H}$  encodes the association uncertainty in the following manner: For measurement  $z_j$  associated with face  $a_j$ , the columns in  $\mathcal{H}$  have zero values except

the ones corresponding to the face  $a_j$ , for the rows corresponding to the measurement. This ensures that only the associated face contributes to the measurement.

- (iii) For measurement  $z_j$  associated with face  $a_j$ , the columns corresponding to the associated face have values  $I - \Lambda_{ij}$  and  $\Lambda_{ij}$ , where  $\Lambda_{ij} = \lambda_{ij}I$  for a scalar  $\lambda_{ij} \in [0, 1]$ .  $\lambda_{ij}$  corresponds to the  $j^{\text{th}}$  measurement from face  $a_j = i$  and is used to denote the exact location of the measurement source along the face.

The total measurement set  $Z^k$  at time step  $k$  can be written in terms of the state  $\mathbf{x}$  and an appropriate measurement source matrix  $\mathcal{H}$  for time step  $k$  as :

$$Z^k = \mathcal{H}\mathbf{x} + w, \quad w \sim \mathcal{N}(0, \mathcal{R}). \quad (2.12)$$

where  $w$  is the stacked noise vector for the measurement set, and  $\mathcal{R}$  is the block diagonal matrix whose elements are the measurement noise covariance for a single measurement  $R$ . The number of block diagonal elements  $R$  in the matrix  $\mathcal{R}$  is equal to the number of measurements in  $Z^k$ . This notation is used in the further sections such that  $\mathcal{R}$  is compatible with the number of measurements being addressed.

It must be noted that  $\mathcal{H}$  in Equation (4.9) is unknown. An example of the measurement source matrix is given in (2.13), where measurements  $z_{11}, z_{12}$  occur from face  $(p_1, p_2)$ ,  $z_{31}$  occurs from face  $(p_3, p_4)$ ,  $z_{41}, z_{42}, z_{43}$  occur from face  $(p_4, p_5)$ , and no measurements occur from the remaining faces. Figure (2.3) shows the Equation (2.13) visually.

$$\begin{bmatrix} z_{11} \\ z_{12} \\ z_{31} \\ z_{41} \\ z_{42} \\ z_{43} \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & I - \Lambda_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Lambda_{12} & I - \Lambda_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Lambda_{31} & I - \Lambda_{31} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Lambda_{41} & I - \Lambda_{41} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Lambda_{42} & I - \Lambda_{42} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Lambda_{43} & I - \Lambda_{43} & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ v \end{bmatrix} + \begin{bmatrix} w_{11} \\ w_{12} \\ w_{31} \\ w_{41} \\ w_{42} \\ w_{43} \end{bmatrix} \quad (2.13)$$

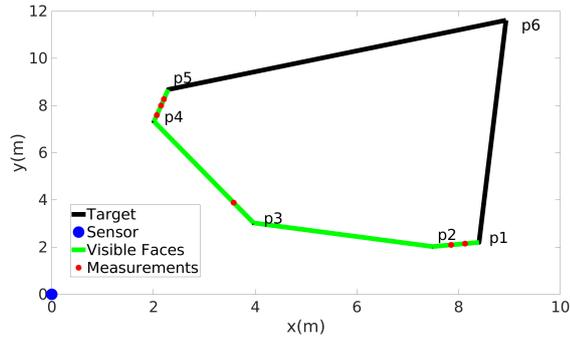


FIGURE 2.3: Visual representation of Equation (2.13)

Treating  $\mathcal{H}$  as a random variable, the posterior conditioned on  $\mathcal{H}$  can be written as :

$$p(\mathbf{x}|Z, \mathcal{H}) \propto p(Z, \mathcal{H}|\mathbf{x})p(\mathbf{x}) \quad (2.14)$$

$$\propto p(Z|\mathcal{H}, \mathbf{x})p(\mathcal{H}|\mathbf{x})p(\mathbf{x}), \quad (2.15)$$

where  $p(Z|\mathcal{H}, \mathbf{x})$  is the likelihood of the measurements  $Z$  conditioned on  $\mathcal{H}$ .

Using (2.13), the likelihood of the measurements  $Z$ , given the measurement source matrix  $\mathcal{H}$  and the state  $\mathbf{x}$ , can be approximated using Gaussian assumption. It must be noted that  $\mathcal{H}$  is kept fixed and its uncertainty is not incorporated

directly into the likelihood. Since the  $\mathcal{H}$  is unknown, the measurement likelihood is one of the measures of consistency of the measurement model.

$$p(Z|\mathcal{H}, \mathbf{x}) = \mathcal{N}(Z; \mathcal{H}\mathbf{x}, \mathcal{R}) \quad (2.16)$$

The probability of a certain measurement source matrix  $\mathcal{H}$ , given the state  $\mathbf{x}$ ,  $p(\mathcal{H}|\mathbf{x})$  is approximated by the probability of the number of measurements as specified by  $\mathcal{H}$  occurring from the faces of  $\mathbf{x}$ ,

$$p(\mathcal{H}|\mathbf{x}) = \prod_i^{N_s} Poisson(n_f^i; N_f^i), \quad (2.17)$$

where  $n_f^i$  is the number of measurements associated to the face  $i$  as specified by  $\mathcal{H}$  and  $N_f^i$  is the average number of measurements generated from the face as calculated using Equation (2.7). If a certain face of the target is not visible but the estimated face corresponding to it is visible, no measurements may be associated to the specified face leading to a low probability of the face. This problem is handled in Section 2.5.5. On the other hand, if  $\mathcal{H}$  associates any measurements to a face of  $\mathbf{x}$  that is not visible to the sensor, then  $p(\mathcal{H}|\mathbf{x}) = 0$ .

The prior for time step  $k$  is assumed to be Gaussian. The predicted density is given below as the prior. The predicted density is calculated using transition dynamics and the updated estimate from the previous time step.

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; F\hat{\mathbf{x}}(k-1|k-1), FP(k-1|k-1)F^T + \Gamma Q \Gamma^T) \quad (2.18)$$

The predicted mean and the covariance for Equation (2.18) are calculated using Equation (3.4).

The optimal method would be to jointly estimate the posterior state  $\mathbf{x}$  and current measurement source matrix  $\mathcal{H}$  to maximize  $p(\mathbf{x}, \mathcal{H}|Z)$  (such as using a grid method). The joint estimation requires the calculation of the posterior  $p(\mathbf{x}, \mathcal{H}|Z)$  for all possible values of  $\mathbf{x}$  and  $\mathcal{H}$  and choosing the values that maximize the posterior. Since the association to the source locations and the number of vertices in the state estimate are unknown, the posterior needs to be evaluated for different combinations of the number of vertices of the state estimate and source locations. This is computationally expensive, especially since the posterior may not have an analytical form. Another method is to estimate  $\mathcal{H}$  and  $\mathbf{x}$ , alternately assuming that one of them is known at each step. It is then possible to obtain an expression for the conditional probability densities  $p(\mathbf{x}|Z, \mathcal{H})$  and  $p(\mathcal{H}|\mathbf{x})$ . The latter method is used (and a framework is developed) to iteratively estimate  $\mathbf{x}$  and  $\mathcal{H}$ , assuming one of them is fixed at a time.

## **2.4 Extended Target Tracking in Multitarget Tracking Framework**

A typical multitarget tracking (MTT) framework for point targets involves :

1. Gating - Measurements are validated using the appropriate confidence gates for existing tracks. New tracks are to be initialized for measurements not falling within the gates of any existing tracks. The confidence gates are constructed using the predicted density of the tracks. Multitarget data association is used if measurements fall into overlapping gates of multiple targets.

2. Association - Various data association algorithms can be used to associate valid measurements to existing tracks. Nearest neighbour association is a simple data association technique and can be used directly in case of no clutter. Some data association techniques that support clutter include Joint Probabilistic Data Association (JPDA) [18],[19] and Multiple Hypotheses Tracker (MHT) [20]. The same measurement may be (probabilistically) associated to multiple targets.
3. Filtering - The existing tracks are filtered using the measurements associated to them. Kalman Filter is used in the case of linear Gaussian models. Extended Kalman Filter can be used to filter targets with nonlinear dynamics and/or nonlinear measurement function by linearization at the estimate [21]. Unscented Kalman Filter (UKF) is yet another method to handle the nonlinearity.[22]. In the case of non-Gaussian noise models, the Particle Filter approach is used, particularly if the noise is multi-modal. The Particle Filter is versatile and can be used in the case of nonlinear models as well [23].
4. Initialization - Measurements not associated to any existing track are used to create new tracks.
5. Track Management - Quality or logic based track management can be used to upgrade tentative tracks to confirmed tracks and delete low-quality tracks [16].

The point multitarget tracking framework can be modified accordingly to handle data streams (measurements) from sources other than point targets as well. The framework can handle data from unknown and possibly time-varying number of sources with association uncertainty, i.e., the target source for each individual measurement is unknown, assuming that the source dynamics and measurement

generation mechanism are known. For an extended target, the different faces/parts of the target can be considered to be the different sources with its dynamics constrained by the fact that it is a rigid body. Since the target now has a finite extent (as opposed to the zero extent of a point target), visibility issues also need to be considered. The framework is described as follows:

1. Association - Measurements are to be associated with visible parts of the target or clutter. Traditional association techniques may need to be modified since the measurement function may be unknown at this point. The nearest neighbour approach can be used directly to associate the measurements to the closest visible face/part of the target.
2. Determination of the measurement function - Using the associated measurements from the previous step, a measurement function has to be determined for each face, and hence for the whole target. The projection of the measurements on the estimated shape can be used to determine the measurement function. The distribution of the measurements along the face can also be accounted for in the measurement function.
3. Gating - It is determined whether the measurements associated with each face fall within the gate of the current face to which they are associated. The measurements which fall outside the gate are to be used for initializing new faces.
4. Filtering - Measurements associated with the existing faces of the target are used to update the target shape and kinematics. For example, a Kalman filter can be used for linear models with Gaussian noise.

5. Initialization - New faces are initialized using the measurements, which do not fall within the gate of any face. New faces are added at the vertices of the existing target shape. Unassociated measurements can be grouped into sets, each of which is used to initialize a new face.
6. Face Management - Faces are either maintained for the next step or deleted, depending on the quality of the face. The quality of the face can be determined using the actual number of measurements associated with the face versus the expected number of measurements occurring from the face.
7. Additional Constraints - Additional constraints (such as the convex constraint) can be applied to restrict the target shape if the constraints are not already incorporated into the filtering step.

This process is continued for each frame until the maximum number of iterations is reached or an optimality criterion is met. The flowcharts for a single iteration of the multitarget tracking framework and the proposed framework are shown in Figures (2.4a),(2.4b).

## **2.5 Implementation**

Assuming a 2D convex target shape, NCV model dynamics, and the absence of clutter, an algorithm is described under the above framework for extended target tracking. Some of the functions used to determine the measurement function are described first.

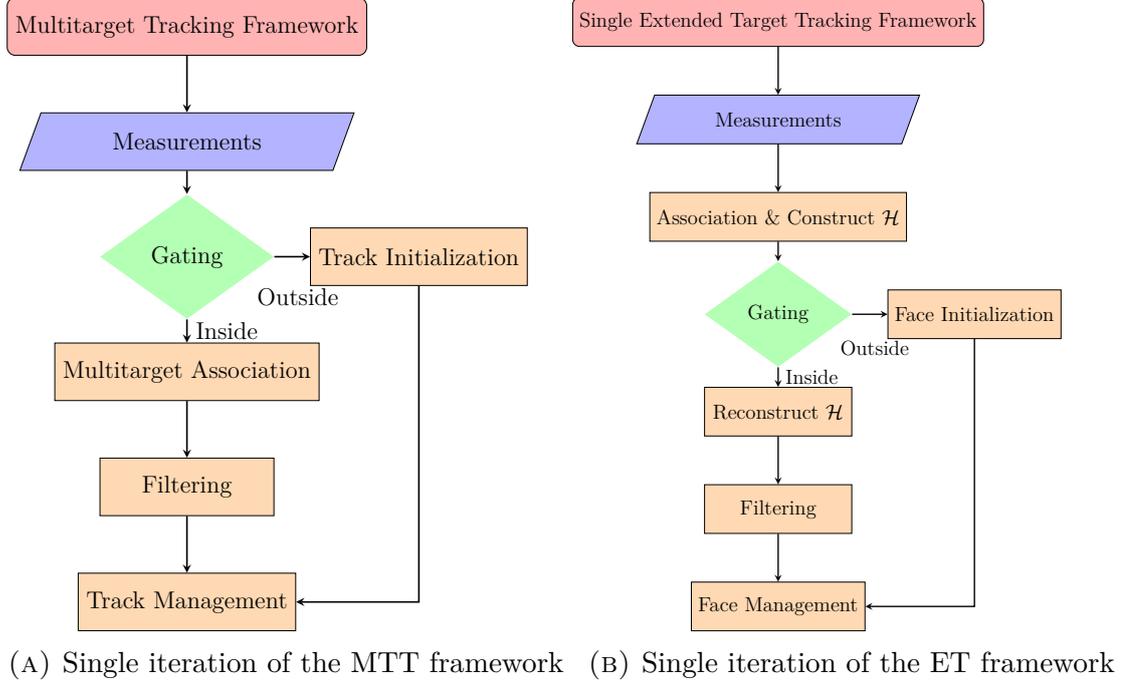


FIGURE 2.4: Flowchart of the frameworks

### 2.5.1 Determining the measurement function - Association of measurement to source location

The measurements are assumed to occur only from the surface of the target with a known measurement noise covariance  $R$ . A linear model is used to represent the measurement function for each measurement from the face with vertices  $p_i$  and  $p_{i^*}$  as follows :

$$z_{ij} = (1 - \lambda_{ij})p_i + \lambda_{ij}p_{i^*} + w_{ij}, \quad w_{ij} \sim \mathcal{N}(0, R). \quad (2.19)$$

The objective is to associate each measurement with a face  $(p_i, p_{i^*})$  and estimate the source location along the face specified by  $\lambda_{ij}$ . The source locations can then

be assembled similar to Equation (2.13) to form the measurement source matrix. Construction of the measurement source matrix is described below:

- (i) Nearest Neighbour association - The measurements are associated to the nearest visible face of the target. It must be noted that this is a greedy association model. The Euclidean distance between a measurement and a face is given by the distance between the measurement and the projection of the measurement on the line along which the face lies. If the projection is outside the limits of the face, the Euclidean distance is the distance of the measurement to the closest vertex of the face. The measurements are grouped by the face to which they are associated.

If the groups of measurements have at least two measurements per set, the following procedure is used-

- (ii) Face estimation - A line/plane is fitted to the face with the set of associated measurements using linear regression. The measurements are assumed to occur from the same (associated) face with zero mean Gaussian errors and a covariance  $R$ . Assumptions about measurement noise need not hold in the case of an incorrect measurement-to-source-location association. A set of measurements associated with the face  $i$  termed as  $Z_i$  can be obtained using the association described previously. The set  $Z_i = \{z_{i1}, z_{i2}, \dots, z_{in_f}\}$ , with each measurement  $z_{ij} = \begin{bmatrix} z_{ij}^x & z_{ij}^y \end{bmatrix}^T$ , where  $z_{ij}^x, z_{ij}^y$  are the  $x$  and  $y$  co-ordinates of the measurement respectively. The objective is to estimate the line along which the source face of the measurements lies. The line is estimated in the slope-intercept form with slope  $m$  and intercept  $c$  as follows -

- (a) A stacked measurement equation is developed using the slope-intercept form of the line for each associated measurement, i.e.,  $z_{ij}^y = mz_{ij}^x + c + w_{ij}^{xy}$ , where  $w_{ij}^{xy}$  is the zero mean Gaussian noise with covariance  $\sigma_x^2 + \sigma_y^2$ . The stacked measurement equation can be written as :

$$Y_i = \begin{bmatrix} X_i & \mathbf{1} \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} + W_i, \quad (2.20)$$

with each element of  $W_i$  being  $w_{ij}^{xy} \sim \mathcal{N}(0, \sigma_x^2 + \sigma_y^2)$ ,

$$\text{and } Y_i = \begin{bmatrix} z_{i1}^y \\ z_{i2}^y \\ \vdots \\ z_{in_f}^y \end{bmatrix}, \quad X_i = \begin{bmatrix} z_{i1}^x \\ z_{i2}^x \\ \vdots \\ z_{in_f}^x \end{bmatrix}, \quad W_i = \begin{bmatrix} w_{i1}^{xy} \\ w_{i2}^{xy} \\ \vdots \\ w_{in_f}^{xy} \end{bmatrix}, \quad \mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}. \quad (2.21)$$

- (b) The estimate of  $\begin{bmatrix} m & c \end{bmatrix}^T$  and its covariance  $P_{mc}$  are then calculated as :

$$\begin{bmatrix} m \\ c \end{bmatrix} = P_{mc} \begin{bmatrix} X_i^T \\ \mathbf{1}^T \end{bmatrix} (\sigma_x^2 + \sigma_y^2)^{-1} Y_i, \quad P_{mc} = \left( \begin{bmatrix} X_i^T \\ \mathbf{1}^T \end{bmatrix} (\sigma_x^2 + \sigma_y^2)^{-1} \begin{bmatrix} X_i & \mathbf{1} \end{bmatrix} \right)^{-1} \quad (2.22)$$

The covariance  $P_{mc}$  is a measure of the consistency of the association. A large covariance  $P_{mc}$  for the line may indicate that the measurements do not belong to the same face.

- (iii) Estimating the exact source location on the face -

- Directly using the projections - Once a set of measurements is associated to a face, the source location of the measurements along the vertices needs to be determined. The source locations can be determined using the distribution of the measurements along the face. The projections of the measurements along the face are used as empirical samples to fit the known distribution. The vertices of the face are calculated using the parameters of the distribution. The source locations of the measurements are then the projections of the measurements along the vertices calculated in the previous step.

The source location between the vertices of the face can also be estimated directly without knowing the distribution of the measurements along the face. The latter method is used in this work. The direct estimation of the vertices using the projections ignores the spatial distributions of the measurements along the face and is not optimal for faces with a skewed distribution of measurements. Using the uniform distribution model in such cases leads to poor estimates for the vertices, which affects the measurement source matrix estimation and in turn the entire estimation process. Additionally, the Face Management step is affected by the wrong choice of the distribution as well.

To estimate the source locations directly, the projections of the measurements onto the estimated line/face are calculated. The vertices are set to be the maximum and minimum of the projections along the line/face. The source locations are calculated as -

$$\hat{\lambda} = \frac{\lambda_z - \lambda_{min}}{\lambda_{max} - \lambda_{min}}, \quad (2.23)$$

where  $\lambda_z$  is the projection of the measurement onto the face,  $\lambda_{min}, \lambda_{max}$  are the minimum and maximum of the projections respectively.

It must be noted that the  $\hat{\lambda}$  is different from  $\lambda$  in (2.9). The latter is unknown, and the former has been estimated using the measurements.

If the set of measurements has only one measurement, the source location along the face  $\hat{\lambda}$  is set to 1. In all cases, the maximum and minimum projections are compared with the vertices of the face to which they are associated. The form of  $\hat{\lambda}$  is such that the minimum projection corresponds to  $p_i$  in Equation (2.19). If the minimum projection is closer to  $p_{i^*}$  instead, then the source locations are changed as  $\hat{\lambda} = 1 - \hat{\lambda}$ . The measurement source matrix  $\mathcal{H}$  is calculated using the source locations.

## 2.5.2 Convex hull and open convex hull

The algorithm uses the convex hull constraint to maintain the convex shape of the target since the shape after filtering may not be convex. A convex set  $\mathcal{K}$  in  $\mathbb{R}^n$  is a set such that

$$\forall p, q \in \mathcal{K}, \quad \gamma p + (1 - \gamma)q \in \mathcal{K}, \quad \gamma \in [0, 1]. \quad (2.24)$$

The convex hull of a set of points in  $\mathbb{R}^n$  is the smallest convex set that contains these points. Such a convex hull is, in fact, a polytope. If the points are random, the resulting convex hull is a random polytope. Algorithms to find the convex hull

of a set of points are described in [24]. The implementation of such algorithms is a part of many libraries [25],[26].

An 'open convex hull'(polytope) is described to account for the uncertainty in the shape. An indicator variable  $\delta_o$  is used for each face. The indicator variable  $\delta_o = 1$  implies that the face is a part of the 'open convex hull'. Meanwhile,  $\delta_o = 0$  implies that the vertices of the face are a part of the convex hull, but the face joining the vertices is not part of the boundary of the target This is used to keep a record of the faces that were estimated using measurements and faces that are only an extension between the vertices of existing faces.

The inputs to the convex hull are either measurements or estimates, both of which are random. The algorithms for the convex hull do not incorporate noise or error in the points, and hence the output can be quite unstable. Vertices of the convex hull output with adjacent faces that have angles greater than  $\theta_{CH}$  are deleted from the convex hull to ensure a more stable output.

### **2.5.3 Initialization of velocity and shape**

Using the measurements  $Z(0)$  from the first frame, the shape  $X_S(0)$  and its covariance  $P_S(0)$  is initialized iteratively as follows :

- (i) All the measurements are initially assumed to occur from a single face and hence are grouped into one single set  $Z_i$ .
- (ii) The measurement source matrix  $\mathcal{H}$  is determined from the measurements.

- (iii) The vertices of the face  $i$  termed as  $f_i$ , and their covariance  $P_i$  are estimated as follows :

$$P_i = (\mathcal{H}^T \mathcal{R}^{-1} \mathcal{H})^{-1}, \quad f_i = \begin{bmatrix} p_i \\ p_{i^*} \end{bmatrix} = P_i \mathcal{H}^T Z_i. \quad (2.25)$$

Note that for the first iteration, a single face is estimated, i.e.,  $i = 1$ . In the following iterations Equation (2.25) can be used to initialize faces  $i = 1$  to  $N_S$ .

- (iv) The following procedure is continued until there is no change in the association of measurements or the maximum number of iterations is reached.

- (a) The measurement source matrix  $\mathcal{H}$  is constructed using the estimate from the previous iteration  $X_S(0), P_S(0)$  (only a single face in the first iteration). The measurements are validated using the face estimates to which they are associated. The measurements  $Z_i = \{z_{ij}\}_{j=1 \text{ to } n_j^i}$  are associated to the face  $i$ , i.e.,  $(p_i, p_{i^*})$  and the covariance of the face is  $P_i$ . The part of the measurement matrix corresponding to the face-measurement combination is  $\mathcal{H}_{ij}$ . A pre-defined threshold  $\Gamma_G$  is a tunable parameter used to determine the gating of the confidence region.

$$\left( z_{ij} - \mathcal{H}_{ij} \begin{bmatrix} p_i \\ p_{i^*} \end{bmatrix} \right)^T S_{ij}^{-1} \left( z_{ij} - \mathcal{H}_{ij} \begin{bmatrix} p_i \\ p_{i^*} \end{bmatrix} \right) \leq \Gamma_G, \quad (2.26)$$

where  $S_{ij} = \mathcal{H}_{ij} P_i \mathcal{H}_{ij}^T + R$ .

In [21] , Chapter 10.4 , the process noise covariance is artificially increased to compensate for the uncertainty due to the linearization of the process and measurement models at the estimate. In the current case, the measurement source matrix  $\mathcal{H}$  is not known accurately. An innovation inflation factor  $\tau_I$  is thus incorporated to compensate for the uncertainty in the measurement model

$$S_{ij} = \mathcal{H}_{ij}P_i\mathcal{H}_{ij}^T + R + \tau_I I_I, \quad \text{where } I_I \text{ is of compatible size as } S_{ij}. \quad (2.27)$$

- (b) The measurements falling inside the validation region are grouped into sets based on the faces to which they are associated. On the other hand, the measurements falling outside the validation region are grouped by the closest visible vertex in the estimate. Together, these sets of measurements are termed as  $Z_T$  consisting of sets  $Z_i$  for faces  $i = 1$  to  $N_S$ .
- (c) All the sets of measurements are initialized separately to form new faces  $f_i$  using (2.25).
- (d) The convex hull of the new faces that are initialized gives the shape estimate. The covariance matrix is changed accordingly. It must be noted that the cross-covariance between the faces is zero since they were initialized separately.
- (v) An array  $\Delta_o(0)$  of indicator variables  $\delta_o$  is created to represent an 'open convex hull'. The measurements are re-associated to the final shape estimate after the iterations are completed. For the face  $(p_i, p_{i^*})$  associated with at

least one measurement, the corresponding  $\delta_o$  is set to 1. For the faces not associated with any measurements, the corresponding  $\delta_o$  is set to 0.

The velocity is initialized according to the specific dynamics model. For example, for the NCV model, the velocity is initialized as  $v(0) = \begin{bmatrix} 0 & 0 \end{bmatrix}$  with an appropriate covariance  $P_v(0)$ . The total initial estimate is -

$$\hat{\mathbf{x}}(0|0) = \begin{bmatrix} X_S(0) \\ v(0) \end{bmatrix}, \quad P(0|0) = \begin{bmatrix} P_S(0) & \\ & P_v(0) \end{bmatrix}. \quad (2.28)$$

---

**Algorithm 1** Shape Initialization

---

```

procedure INITIALIZE_SHAPE( $Z(0)$ , iter_max)
   $\mathcal{H} \leftarrow$  construct_measurement_source_matrix( $Z(0)$ ) (associated to single face)
   $P_S(0) \leftarrow (\mathcal{H}R^{-1}\mathcal{H})^{-1}$    $X_S(0) \leftarrow P_S(0)\mathcal{H}^T R^{-1}Z(0)$ 
  for r=1, iter < iter_max do
     $\mathcal{H} \leftarrow$  construct_measurement_source_matrix( $X_S(0)$ ,  $Z(0)$ )
     $Z_T \leftarrow$  validate_measurements( $X_S(0)$ ,  $P_S(0)$ ,  $\mathcal{H}$ ,  $Z(0)$ )
    for  $Z_i \in Z_T$  do
       $\mathcal{H} \leftarrow$  construct_measurement_source_matrix( $Z_i$ )
       $P_i \leftarrow (\mathcal{H}^T R^{-1}\mathcal{H})^{-1}$    $f_i \leftarrow P_i\mathcal{H}^T R^{-1}Z_i$ 
       $X_S(0) \leftarrow [X_S(0) \ f_i]$    $P_S(0) \leftarrow$  block_diagonal( $P_S(0)$ ,  $P_i$ )
     $X_S(0)$ ,  $P_S(0) \leftarrow$  convex_hull( $X_S(0)$ ,  $P_S(0)$ )
   $\Delta_o(0) \leftarrow$  face_measurement_association( $X_S(0)$ ,  $Z(0)$ )

```

---

## 2.5.4 Prediction

The state and covariance are predicted across time steps  $k$  using the dynamics equation, Equation (3.4) while utilizing the fact that the target is a rigid body.

The array of indicator variables remains unchanged.

$$\hat{\mathbf{x}}(k+1|k) = F(k)\hat{\mathbf{x}}(k|k), P(k+1|k) = F(k)P(k|k)F(k)^T + \Gamma(k)Q(k)\Gamma(k)^T \quad (2.29)$$

$$\Delta_o(k+1|k) = \Delta_o(k|k) \quad (2.30)$$

### 2.5.5 Update

Given the measurements  $Z(k) = \left[ z_1^T(k) \ z_2^T(k) \ z_3^T(k) \ \dots \ z_{N_k}^T(k) \right]^T$  and the estimate predictions  $\hat{\mathbf{x}}(k+1|k), P(k+1|k), \Delta_o(k+1|k)$ , the objective is to estimate the updated state  $\hat{\mathbf{x}}(k+1|k+1), P(k+1|k+1), \Delta_o(k+1|k+1)$ . The following operations are performed until the maximum number of iterations is reached and the updated estimate with the minimum cost is chosen. The cost metric  $\Delta_{cost}$  is the distance error between the current measurements to their closest point in the estimate.

The variables  $x^r, P^r, \Delta_o^r$  are temporary variables used in the iteration set initially to the predicted estimates  $\hat{\mathbf{x}}(k+1|k), P(k+1|k), \Delta_o(k+1|k)$  respectively.

- (i) The measurement source matrix  $\mathcal{H}$  is constructed using  $x^r$  and the current measurements  $Z(k)$ .
- (ii) The measurements are validated for the face to which they are associated by using Equations (4.58), (2.27) with the  $\mathcal{H}$  constructed in the previous step. The measurements falling outside the validation region are grouped by the nearest visible vertex in the estimate. Measurements associated with the faces

that have  $\delta_o^r = 0$  (the face is an extension between existing vertices and was not initialized using measurements) are treated as new measurements. They are grouped with the measurements falling outside the validation regions to initialize new faces. The measurements inside the validation region of at least one face are termed as  $Z_C$  and the rest of the measurements are termed as  $Z_N$ .

- (iii) For the first iteration, the validation step is skipped and all the measurements are used to construct  $\mathcal{H}$  and update the predicted estimate. For the second iteration, the validation step is performed but all the measurements falling outside the validation region are grouped into a single set. The first two iterations accommodate the scenario where zero or one new face is added for the current frame. The procedure described above is followed for the rest of the iterations.
- (iv) The measurement source matrix  $\mathcal{H}$  is reconstructed using only the measurements falling inside the validation region  $Z_C$ , and the predicted state  $\hat{\mathbf{x}}(k+1|k)$ . Note that this is the unchanged state from before the iterations begin. The state estimate is updated using a Kalman Filter according to the equations given below -

$$S(k+1) = \mathcal{H}(k+1)P(k+1|k)\mathcal{H}(k+1)^T + \mathcal{R} + \tau_I I_I. \quad (2.31)$$

$S(k+1)$  has been modified in a similar way as Equation (2.27) by using the inflation factor  $\tau_I$ . This modification in the innovation is used to compensate for the uncertainty in the unknown measurement source matrix  $\mathcal{H}(k+1)$  and

leads to better association and validation of measurements.

$$W(k+1) = P(k+1|k)\mathcal{H}(k+1)^T S(k+1)^{-1} \quad (2.32)$$

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + W(k+1) \left[ Z_C - \mathcal{H}(\hat{\mathbf{x}}(k+1|k)) \right] \quad (2.33)$$

$$P(k+1|k+1) = P(k+1|k) - W(k+1)S(k+1)W(k+1)^T \quad (2.34)$$

- (v) The measurements  $Z_N$ , which fall outside the validation region, are grouped into sets according to the nearest visible vertex in the estimate. New faces  $F_N = \{f_{1N}, f_{2N}, \dots\}$  with covariances  $P_N = \{P_{1N}, P_{2N}, \dots\}$  are initialized using each set according to Equation (2.25). These faces are initialized independently of the estimate (zero cross-covariance). The new faces are combined (appended) with the estimated shape and the covariance is modified accordingly. The  $\Delta_{oN}$  for the new faces is set to 1, since they have been initialized using the measurements.
- (vi) Face Management - The probability of association of the number of measurements to each visible face is calculated according to (2.17). Faces with probability falling below the threshold  $\Gamma_{FM}$  are deleted. If the estimated face is visible but the corresponding face in the true target is not visible, zero measurements will occur from that face. This situation occurs when the angle subtended by the face at the sensor is small (the face is 'barely' visible). The probability of the number of measurements associated to such a face is low. To avoid deleting such faces, a new condition is introduced. When the angle subtended by the face at the sensor  $\alpha_S$  is less than a threshold angle for Face Management  $\theta_{FM}$ , the face is not deleted.

- (vii) The convex hull constraint is applied to the resulting shape, and the covariance is modified accordingly. Since some of the faces may get deleted, it is necessary to change the  $\Delta_o^r$  appropriately. The projections of the faces with  $\delta_o^r = 1$  onto the constrained shape are determined. The  $\delta_o^r$  is set to 1 for the faces in the constrained shape along which the maximum projections of the previous valid faces lie. Note that if the face is still present in the constrained shape, the maximum projection lies along itself.

Essentially, the estimate from the previous iteration is used to validate the current measurements into 2 sets. The set of measurements falling inside the validation region is used to update the unchanged predicted state. The other set of measurements is used to initialize new faces. The updated state is combined with the new faces and the convex hull operation is performed to generate the estimate for the current iteration. At the end of the iterations, the estimate with the minimum cost is chosen.

Iterative Method of Probability Updating - The state  $\mathbf{x}$  and the measurement function  $\mathcal{H}$  are estimated alternately, keeping one of them fixed at a time.

- (i) Estimating  $\hat{\mathbf{x}}$  while keeping  $\mathcal{H}$  fixed - The posterior probability  $p(\mathbf{x}|Z, \mathcal{H})$  in (2.15) is estimated from the likelihood  $p(Z|\mathcal{H}, \mathbf{x})$  and the prior probability  $p(\mathbf{x})$  by applying the filtering step. The track management step is used to reject any faces with low likelihood  $p(\mathcal{H}|\hat{\mathbf{x}})$ .
- (ii) Estimating  $\mathcal{H}$  while keeping  $\hat{\mathbf{x}}$  fixed - The estimated state  $\hat{\mathbf{x}}$  is used to construct  $\mathcal{H}$  for the next iteration.

---

**Algorithm 2** State Update

---

```

procedure STATE_UPDATE( $\hat{\mathbf{x}}(k+1|k), P(k+1|k), \delta(k+1|k), Z(k), \text{iter\_max}$ )
   $x^r, P^r, \Delta_o^r \leftarrow \hat{\mathbf{x}}(k+1|k), P(k+1|k), \delta_o(k+1|k)$     $\mathbb{X}, \mathbb{P}, \mathbb{D} \leftarrow [], [], []$ 
  for  $r = 1 : \text{iter\_max}$  do
     $\mathcal{H} \leftarrow \text{construct\_measurement\_source\_matrix}(x^r, P^r, Z(k))$ 
     $Z_C, Z_N \leftarrow \text{validate\_measurements}(x^r, P^r, \Delta_o^r, \mathcal{H})$ 
     $F_N, P_N, \Delta_{oN} \leftarrow \text{initialize\_new\_faces}(Z_N)$ 
     $\mathcal{H} \leftarrow \text{construct\_measurement\_source\_matrix}(\hat{\mathbf{x}}(k+1|k), Z_C)$ 
     $x^r, P^r \leftarrow \text{KF\_update}(\hat{\mathbf{x}}(k+1|k), P(k+1|k), \mathcal{H}, Z_C)$ 
    Equations (2.31) to (2.34)
     $x^r \leftarrow \text{track\_management}(x^r, \Gamma_{FM})$ 
     $x^r, P^r, \Delta_o^r \leftarrow \text{combine\_states}(x^r, P^r, \Delta_o^r, F_N, P_N, \Delta_{oN})$ 
     $x^r, P^r, \Delta_o^r \leftarrow \text{convex\_hull\_constraint}(x^r, P^r, \Delta_o^r)$ 
     $\mathbb{X}, \mathbb{P}, \mathbb{D} \leftarrow [\mathbb{X}, x^r], [\mathbb{P}, P^r], [\mathbb{D}, \Delta_o^r]$ 
   $\hat{\mathbf{x}}(k+1|k+1), P(k+1|k+1), \delta(k+1|k+1) \leftarrow \text{min\_cost\_estimate}(\mathbb{X}, \mathbb{P}, \mathbb{D})$ 

```

---

## 2.6 Results

Simulations are performed to test the performance of the proposed algorithm for various scenarios :

1. Different levels of measurement noise.
2. Different levels of measurements per unit effective area.
3. Targets with different shapes.

In all three cases, the target moves across the sensor's FOV (as shown in Figure (2.5b)) in a manner such that different parts of the target are visible to the sensor at different times.

Three metrics are used to test the performance of the algorithm :

1. Root Mean Square Error (RMSE) of the velocity.
2. RMSE of the center of the target. It must be noted that the true center of the target may be unobservable depending on the visibility and the measurements obtained. Hence, this metric should be interpreted accordingly.
3. Intersection over Union (IoU) - This metric quantifies the similarity between two shapes. It is defined as -

$$\text{IoU} = \frac{\text{Area}(\text{Intersection of the shapes})}{\text{Area}(\text{Union of the shapes})}$$

The optimum value ( in this case, the maximum ) is achieved when both the shapes are the same and  $\text{IoU} = 1$ . Similar to the previous metric, the total shape of the target may be unobservable and hence the metric should be interpreted accordingly.

The scenario chosen for the simulation is such that different parts of the target are visible to the sensor at different times. In the initial time steps, only a part of the target may be visible to the sensor and the whole shape is unobservable as the parts that are not visible do not generate any measurements. Over time, more parts of the target become visible and the whole shape is observable if measurements were obtained from all parts of the target over time. For a 2D convex polytope, if  $N_S - 1$  faces are observable, then the whole shape is observable. The RMSE of the center and IoU results that follow reflect this unobservability. Initially, the RMSE of the center is large and the IoU is low as only a part of the target is observable.

The performance improves over time as more parts of the target become visible.

The performance of the proposed algorithm is compared with the existing algorithm - GP [7] described below.

### 2.6.1 Gaussian Process (GP)

In this approach, the extent of the target is described by a radial function modeled as a GP :

$$f(\theta) \sim \mathcal{GP}(0, \kappa(\theta, \theta') + \sigma_r^2). \quad (2.35)$$

$\theta$  is the input for the radial function  $f(\theta)$ , which describes the radial extent of the target at the given input. The covariance function  $\kappa(\theta, \theta')$  generally used ( and also used in the present work ) is the squared exponential (SE) function :

$$\kappa(\theta, \theta') = \sigma_f^2 \exp\left(-\frac{|\theta - \theta'|}{2l^2}\right). \quad (2.36)$$

$l$  is the length scale and  $\sigma_f^2$  is the prior variance. The process model is given by -

$$\begin{aligned} \mathbf{x}(k+1) &= F(k)\mathbf{x}(k) + v(k), \quad v(k) \sim \mathcal{N}(0, Q(k)) \\ \text{with } \mathbf{x}(k) &= \begin{bmatrix} x_c(k) & \psi(k) & x^f(k) \end{bmatrix}. \end{aligned} \quad (2.37)$$

where  $x_c, \psi, x^f$  describe the center, orientation and the radial function for evenly spaced  $\theta^f$ s in  $[0, 2\pi]$ , respectively. The measurement model for the measurement

$z_{kl}$  obtained at time  $k$  is given as -

$$z_{kl} = h_{kl}(\mathbf{x}(k)) + e_{kl} \quad e_{kl} \sim \mathcal{N}(0, R_{kl}), \quad (2.38)$$

$$h_{kl}(x(k)) = x_c(k) + p_{kl}H^f(\theta_{kl}(x_c(k), \psi(k)))x^f(k), \quad (2.39)$$

$$H^f(\theta_{kl}(x_c(k), \psi(k))) = \mathcal{K}(\theta_{kl}(x_c(k), \psi(k)), \theta^f)\mathcal{K}(\theta^f, \theta^f)^{-1}, \quad (2.40)$$

$$R_{kl} = p_{kl}R_{kl}^f p_{kl}^T + R. \quad (2.41)$$

Using  $\hat{\mathbf{x}}(0) \sim (\mu_0, P_0)$  as the initialization, an EKF is used to recursively update the state.

## 2.6.2 Scenario

The target shape is shown in Figure [2.5a]. The sensor is located at  $\begin{bmatrix} 0m & 0m \end{bmatrix}^T$ . The target is initially centered at  $\begin{bmatrix} 30m & 40m \end{bmatrix}^T$ . The initial velocity of the target is  $\begin{bmatrix} -3ms^{-1} & -4ms^{-1} \end{bmatrix}^T$ . The target trajectory is such that different parts of the target are visible to the sensor at different times. The sampling time  $T$  between frames is 1s. A sample target trajectory is shown in Figure [2.5b].

The measurement noise covariance  $R$  and the power spectral density of the process noise  $q$  are given below.

$$R = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}, \quad q = 0.1m^2s^{-3}.$$

The tracker parameters are as follows :

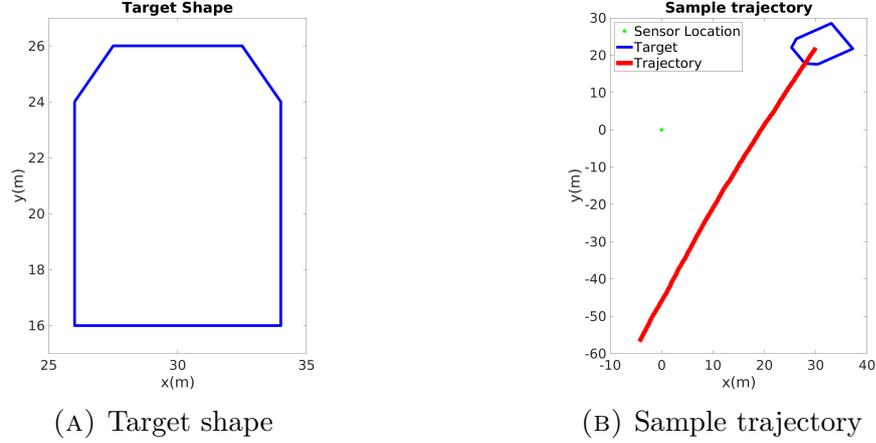


FIGURE 2.5: Target shape and sample trajectory

- (i) Threshold for confidence region of gating  $\Gamma_G = 0.75$ .
- (ii) Threshold for face management  $\Gamma_{FM} = 10^{-9}$ ,  $\theta_{FM} = 10$  degrees.
- (iii) Threshold angle for convex hull  $\theta_{CH} = 175$  degrees.
- (iv) Innovation inflation factor to account for measurement model mismatch -  $\tau_I = 0.1$
- (v) Maximum number of iterations for initialization and update - `iter_max` = 10.

The tracker is initialized with velocity  $\begin{bmatrix} 0 & 0 \end{bmatrix}^T m s^{-1}$  and covariance

$$P_v(0) = \begin{bmatrix} 6^2 & 0 \\ 0 & 6^2 \end{bmatrix} m^2 s^{-2}.$$

The GP parameters are as follows :  $\sigma_f = 2m$ ,  $\sigma_r = 2m$ ,  $l = \frac{\pi}{4}$  and  $\alpha = 0.0001$ . Further, 16 points are used to represent the target contour, i.e., the target shape consists of  $r_{gp}$  for 16 evenly sampled angles  $\theta_{gp}$  between  $[0, 2\pi]$ .

The GP estimate can be self-intersecting and hence the convex hull of the estimate

is used to calculate the results. Note that 50 Monte Carlo runs are used to obtain the results in the present study.

The simulations were performed on MATLAB R2021b on an Intel Core i7-8550U (1.80GHz) processor with 16GB of RAM.

### 2.6.3 Different levels of measurement noise

The algorithms are tested for different values of the measurement noise covariance parameters  $\sigma_x = \sigma_y$  while keeping the number of measurements per unit effective area fixed,  $\rho = 100m^{-1}$ . The values for the standard deviations of the measurement noise covariance used are  $\{0.01m, 0.05m, 0.1m, 0.15m, 0.2m\}$ . Since the GP model does not account for visibility issues, one set of simulations is performed assuming all the faces of the target are visible at all times. However, the number of measurements still depends on the sensor-target geometry.

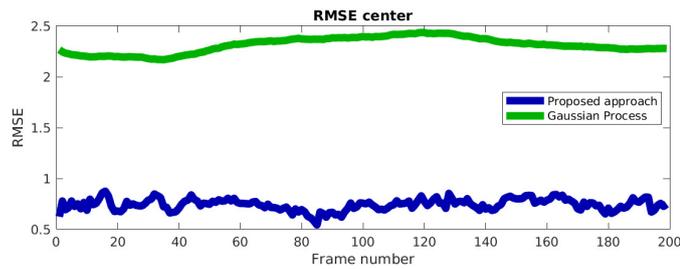


FIGURE 2.6: RMSE center for  $\sigma_x = \sigma_y = 0.1m$  and  $\rho = 100m^{-1}$  when all faces are visible.

Figures (2.6),(2.7),(2.8) show the results when all the faces of the target are visible. Snapshots of the estimates at different times are shown in Figure (2.9)

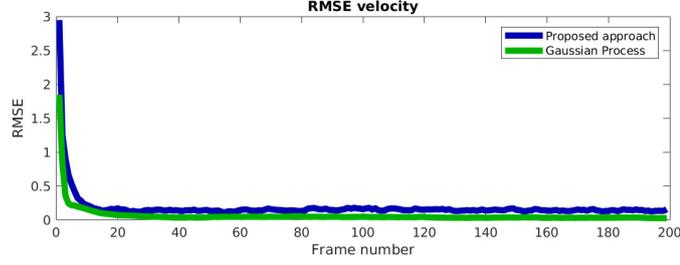


FIGURE 2.7: RMSE velocity for  $\sigma_x = \sigma_y = 0.1m$  and  $\rho = 100m^{-1}$  when all faces are visible.

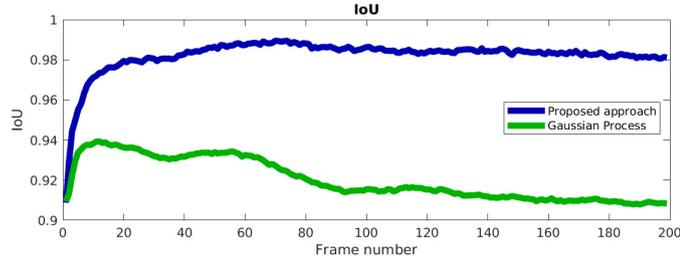
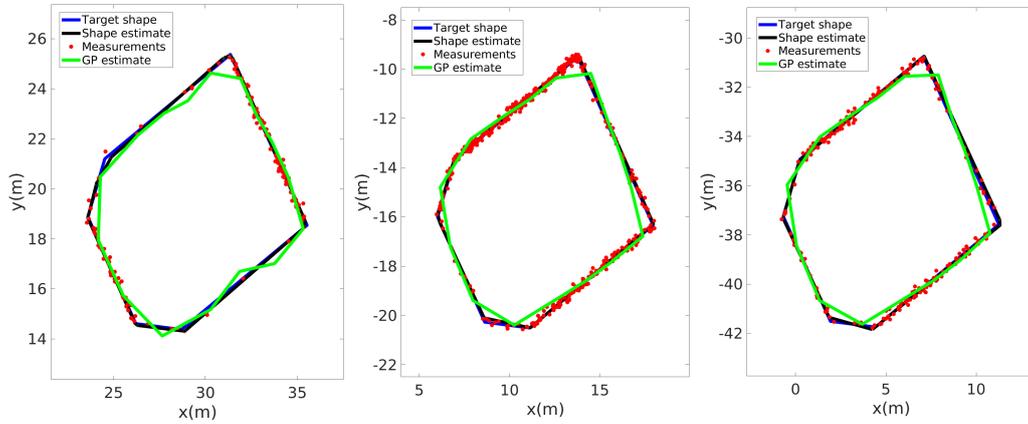


FIGURE 2.8: IoU results for  $\sigma_x = \sigma_y = 0.1m$  and  $\rho = 100m^{-1}$  when all faces are visible.



(A) Frame 10 estimates (B) Frame 100 estimates (C) Frame 150 estimates

FIGURE 2.9: Estimates of the target shape at different times with  $\sigma_x = \sigma_y = 0.1m$  and  $\rho = 100m^{-1}$ . The true target shape is given in blue, and the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green

It can be seen that the results (particularly IoU) are comparable when the visibility issues are not considered. The performance of the proposed algorithm is slightly better since the target shape is a convex polytope, which is more compatible with the proposed model. Importantly, the RMSE of the center for the GP estimate is biased.

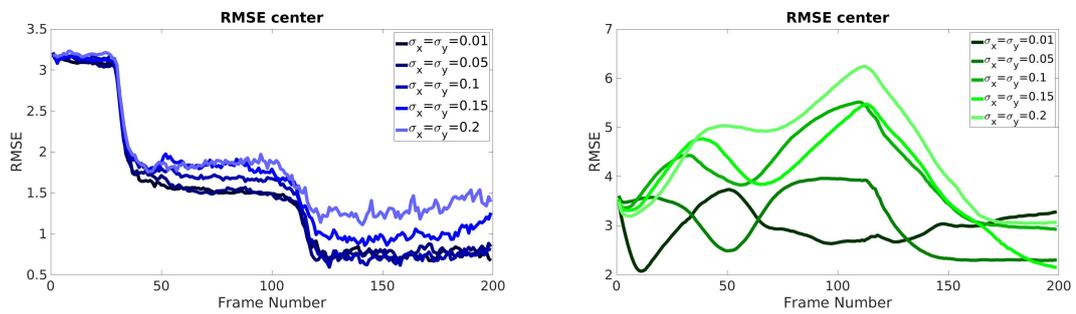


FIGURE 2.10: RMSE center for different levels of measurement noise. Results for the proposed approach are on the left and the GP results are on the right

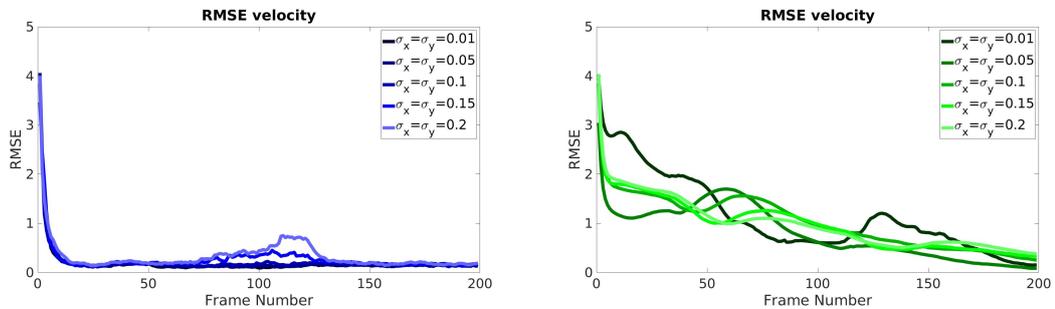


FIGURE 2.11: RMSE velocity for different levels of measurement noise. Results for the proposed approach are on the left and the GP results are on the right

Figures (2.10),(2.11),(2.12) show the results when the visibility issues are considered. Snapshots of the estimates at different times are shown in Figure (2.13)

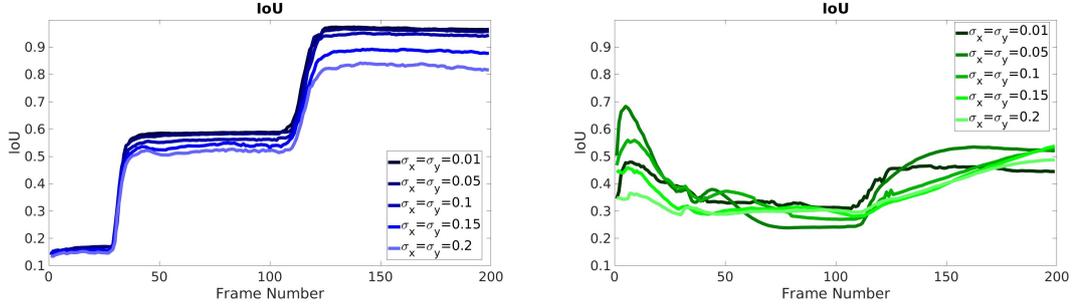


FIGURE 2.12: IoU results for different levels of measurement noise. Results for the proposed approach are on the left and the GP results are on the right

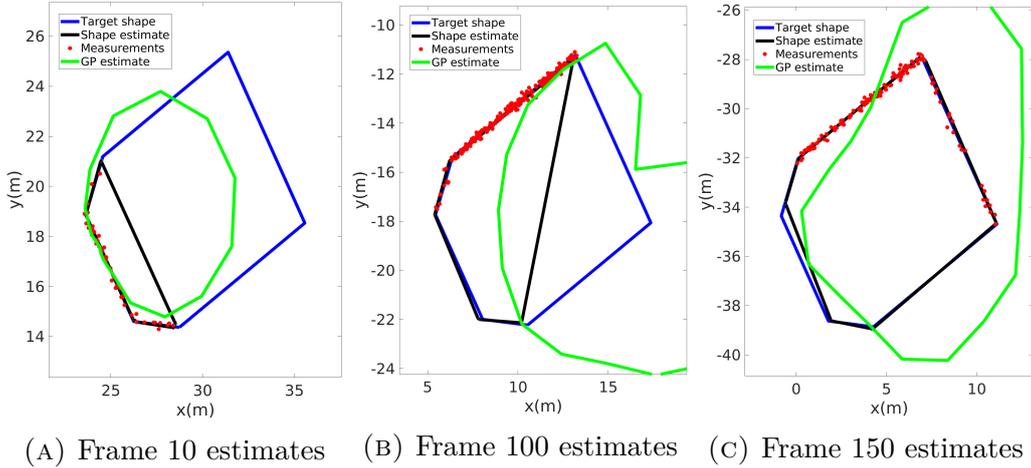


FIGURE 2.13: Estimates of the target shape at different times for  $\sigma_x = \sigma_y = 0.1m$ ,  $\rho = 100m^{-1}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green.

The proposed algorithm outperforms GP when visibility issues are considered. The plots for the estimates of the center and IoU show that the initial values for GP are apparently better than the proposed algorithm. However, from Figure (2.13a), it can be seen that the GP shape estimate is larger than the visible part of the target up to frame 10, which explains the apparent improvement in performance. Hence, the proposed algorithm represents the observable part of the target

better than GP.

From Figures (2.13a),(2.13b),(2.13c), it can be seen that the number of parameters used to represent the target shape changes dynamically as more parts of the target become visible over time.

The effect of change in the measurement noise covariance on the proposed algorithm can be inferred from the error plots (2.10),(2.11),(2.12), i.e., the performance improves with a decrease in the measurement noise covariance. The effect is not as pronounced when  $\sigma_x = \sigma_y$  take values  $\{0.01m, 0.05m, 0.1m\}$ , but as the measurement noise covariance increases to values  $\{0.15m, 0.2m\}$ , the performance degrades gradually. The RMSE of the velocity plot (3.7) is similar for all levels of measurement noise covariance. However, the performance for measurement noise covariance values  $\{0.15m, 0.2m\}$  becomes slightly worse between frames 100 and 120. This is due to the joint uncertainty in the shape and kinematics as more faces of the target become visible over this time. The sharp transitions in the plots show that the estimate becomes closer to the true target state as more parts of the target become visible. The algorithm is able to estimate the true target state with  $N_S - 1$  faces being visible over time (since the target shape is constrained to be convex).

The GP approach is unable to handle the visibility issues, as seen from the snapshots over time (2.13). The estimate worsens over time when only a few faces of the target are visible and improves slightly as more parts of the target become visible.

### 2.6.4 Different levels of average number of measurements

Simulations are performed under the same scenario as described before, however the measurement noise covariances are kept fixed  $\sigma_x = \sigma_y = 0.1m$  and the number of measurements per unit effective area  $\rho$  is varied. The different values of  $\rho$  used are  $\{100m^{-1}, 70m^{-1}, 40m^{-1}, 20m^{-1}\}$ . Figures (2.14),(2.15),(2.16) show the results when the visibility issues are considered. Snapshots of the estimates at different times are shown in Figure (2.17)

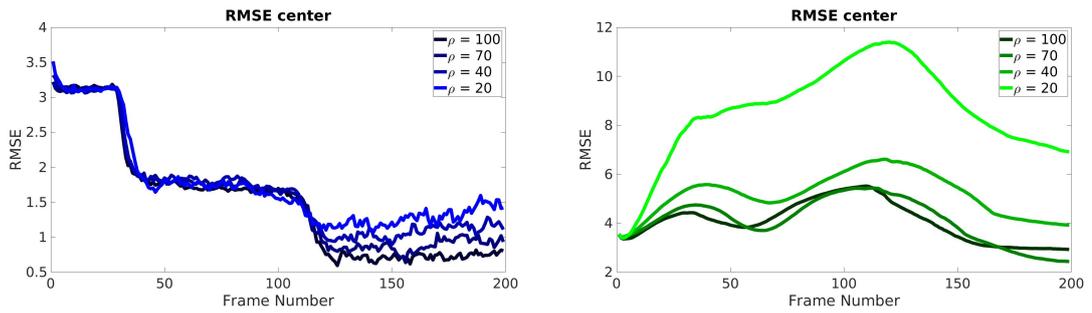


FIGURE 2.14: RMSE center for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right

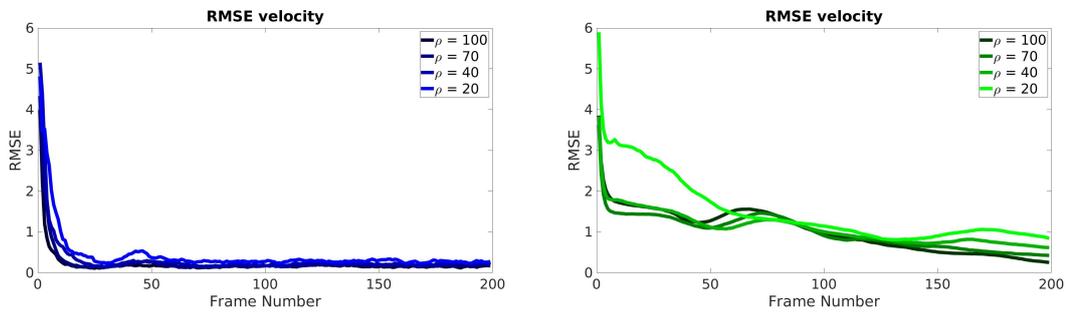


FIGURE 2.15: RMSE velocity for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right

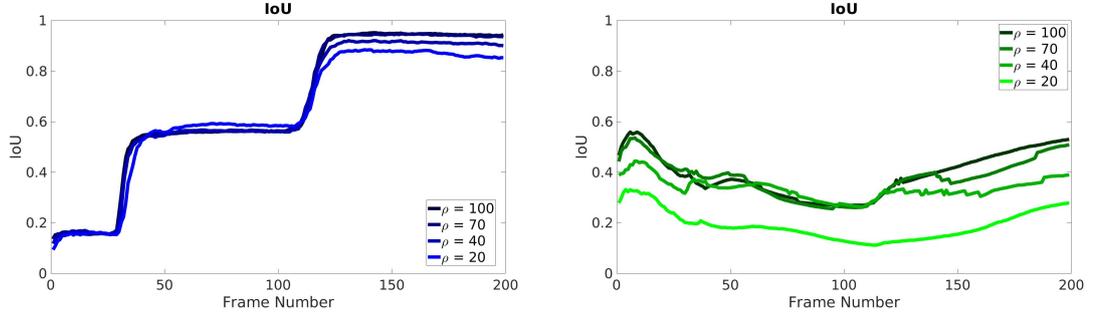


FIGURE 2.16: IoU results for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right

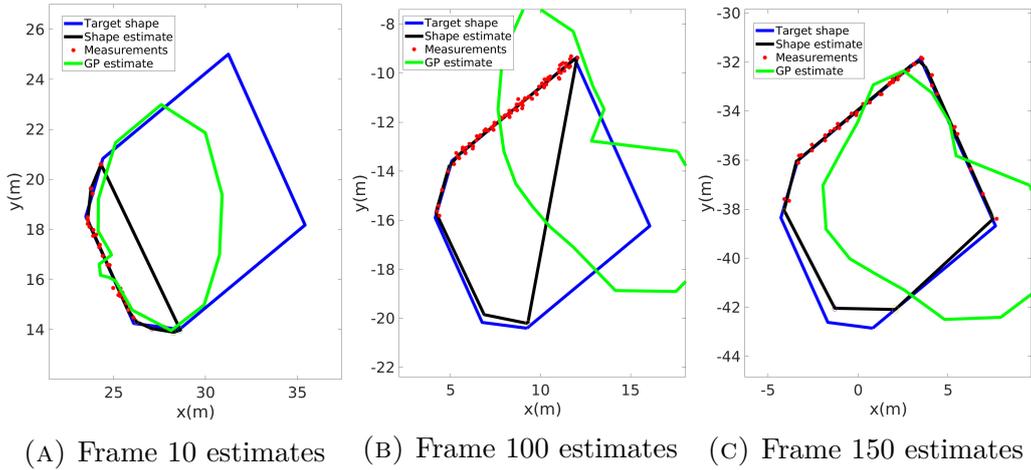


FIGURE 2.17: Estimates of the target shape at different times  $\sigma_x = \sigma_y = 0.1m$ ,  $\rho = 40m^{-1}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green

The proposed algorithm is robust at different levels of measurements from the target as seen from (2.14),(2.15),(2.16). The performance degrades slightly when the level of number of measurements is sufficiently lowered ( $\rho = 20m^{-1}$ ). Notably, the transition in the error plots is slower in case of  $\rho = 20m^{-1}$  compared to other scenarios

The apparent improvement in the performance of GP in the initial time steps has

$\rho$ ( $m^{-1}$ )	Runtime per frame using proposed algorithm for 10 iterations (seconds)	Runtime per frame using GP (seconds)
20	0.1364	0.0173
40	0.1561	0.0301
70	0.1982	0.0542
100	0.2674	0.0973

TABLE 2.1: Table of runtimes for different average levels of measurements

been addressed in Section 2.6.3 - the scenario when visibility issues are considered. Overall, the proposed algorithm is better at representing the observable parts of the target.

The simulation times for the proposed algorithm and GP are listed in Table (2.1) for different levels of average number of measurements. It can be seen that the runtime per frame for the proposed algorithm is significantly greater than GP, however the runtime per iteration is comparable with GP. It must be noted that the code for the proposed algorithm is not optimized for runtime. It can be seen that the difference in runtimes for different average number of measurements is small, i.e., a significant amount of runtime is spent in the setup of the algorithm, which is independent of the number of measurements. Currently, the algorithm uses simple nearest neighbour association and runs till the maximum number of iterations is reached. Better association techniques and an optimality criterion to end the iterations can be developed to minimize the number of iterations, and to have comparable runtime as GP.

## 2.6.5 Performance for different shapes

The scenario is kept the same with  $\sigma_x = \sigma_y = 0.1m$  and  $\rho = 100m^{-1}$ , but the target shape is changed. The target shapes used are rectangle, pentagon and hexagon, as shown in Figure (2.18), and the original custom shape used in the previous scenarios.

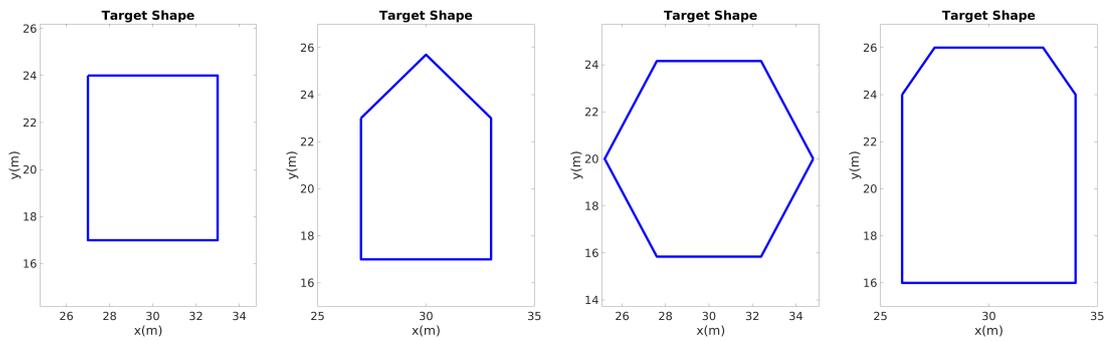


FIGURE 2.18: Rectangle, Pentagon, Hexagon, and custom target shapes

The results are shown in Figures (2.19),(2.20),(2.21); the snapshots of the estimates at different times are shown in (2.22),(2.23),(2.24); and the snapshots with the original shape are shown in (2.13).

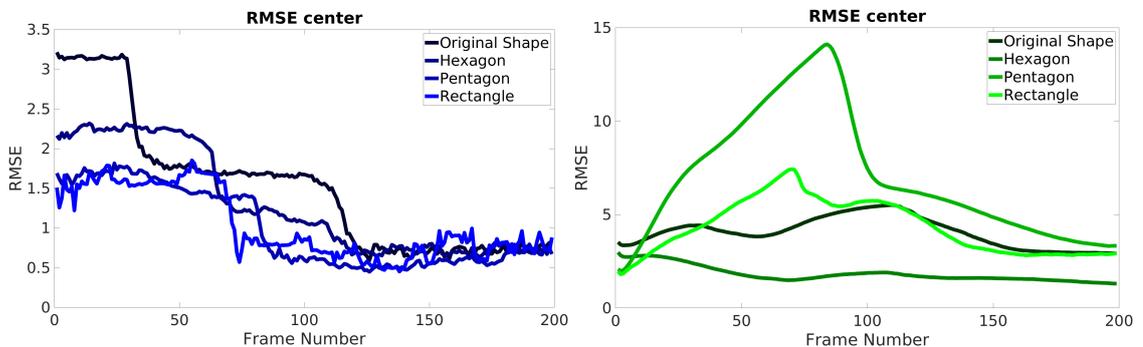


FIGURE 2.19: RMSE center for different shapes. Results for the proposed approach are on the left and the GP results are on the right

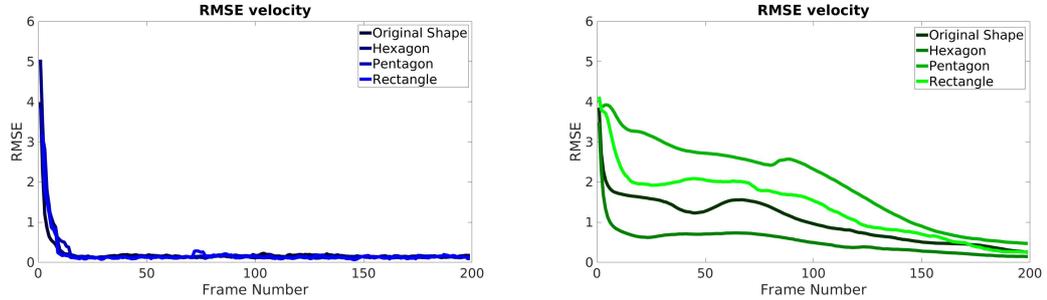


FIGURE 2.20: RMSE velocity for different shapes. Results for the proposed approach are on the left and the GP results are on the right

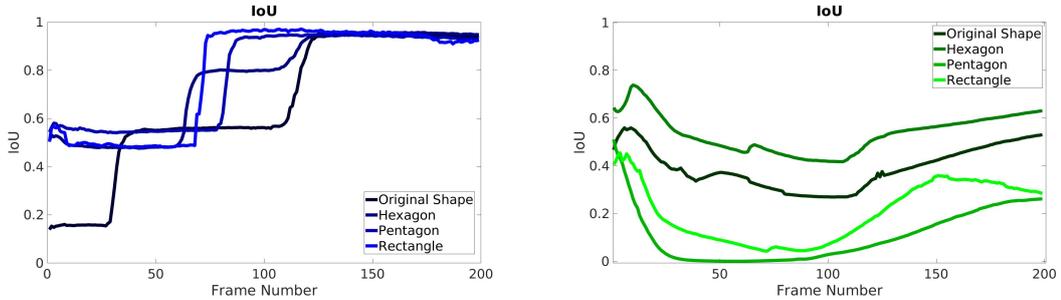


FIGURE 2.21: IoU results for different shapes. Results for the proposed approach are on the left and the GP results are on the right

The algorithm is effectively able to track targets of different shapes, as seen in the results. The visibility of different parts of the target varies with its shape. The algorithm tracks the parts of the target visible up to the current time. It can be clearly seen from Figures (2.19),(2.21) that the performance becomes similar once the 'whole' target becomes visible (approximately frame 75 for rectangle, frame 85 for pentagon, frame 125 for hexagon and frame 120 for the original shape). In fact, only  $N_S - 1$  faces of the target are visible over time and these faces are sufficient to track the entire shape of the target with convex constraint.

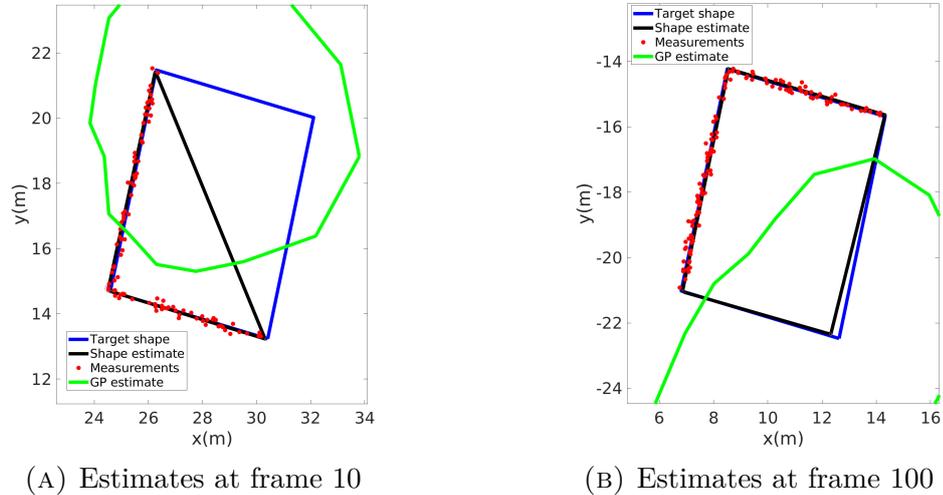


FIGURE 2.22: Estimates of the target shape at different times  $\sigma_x = \sigma_y = 0.1m$ ,  $\rho = 100m^{-1}$  for rectangle target shape. The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green

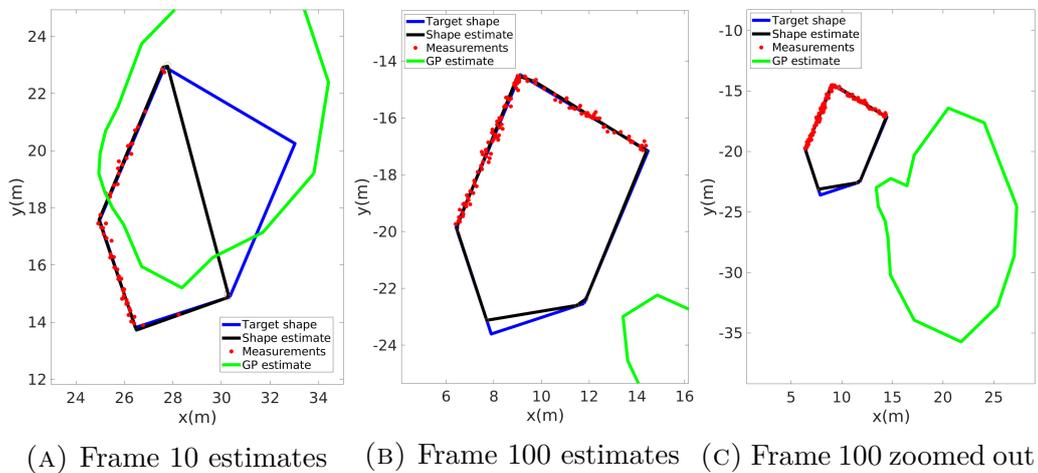


FIGURE 2.23: Estimates of the target shape at different times  $\sigma_x = \sigma_y = 0.1m$ ,  $\rho = 100m^{-1}$  for pentagon target shape. The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green

GP is unable to effectively track the different shapes when visibility issues are considered.

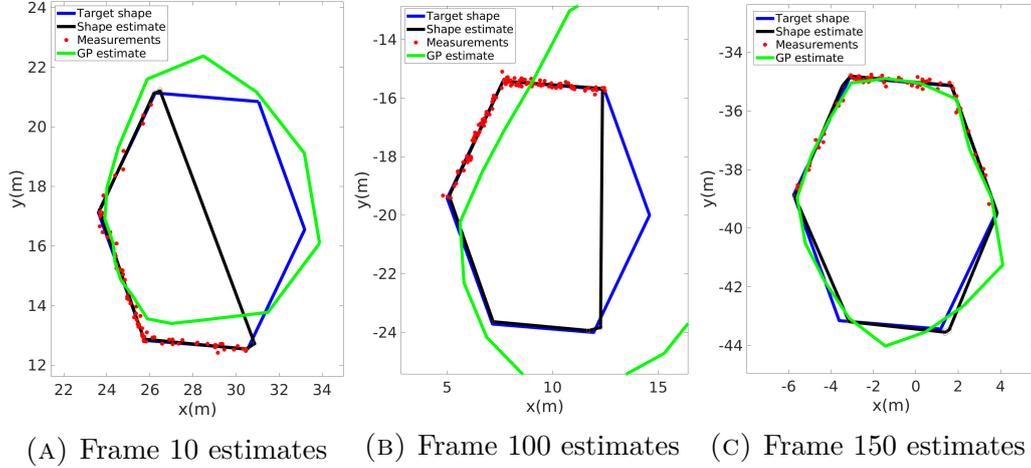


FIGURE 2.24: Estimates of the target shape at different times for  $\sigma_x = \sigma_y = 0.1m$ ,  $\rho = 100m^{-1}$  and hexagon target shape. The true target shape is given in blue, the measurements at the current time are given in red. The estimate using the proposed algorithm is in black and the GP estimate is in green

## 2.6.6 Discussion and Remarks

Some highlights and insights about the results are discussed below. It is seen from the results that both the proposed approach and the Gaussian Process were able to handle the Extended Target Tracking problem when the target is fully visible. However, in reality, the target may only be partially visible at each frame with more parts of the target being visible over different frames. The proposed algorithm is able to handle this problem of self-occlusion whereas the Gaussian Process estimate is unsatisfactory. This is made possible by using the proposed Extended Target Tracking framework and by representing the shape of the target only by its boundary. Since different faces of the Extended Target are treated as separate targets, it is possible to distinguish between the faces that are visible and those that are not visible. Measurements can then be associated to the local visible

faces. The filtering, which maintains the constraint between consecutive faces at the vertex is done jointly. The proposed approach is able to handle only partially visible targets by estimating only the boundary since, in such cases, the actual center of the target is not observable. The shape estimate supports adding new faces as well for the new parts of the target that become visible over frames. It is seen from the RMSE center, RMSE velocity and IoU metrics that the algorithm responds well to changes in the parameters of measurement noise covariance and average number of measurements. The performance of the algorithm varies corresponding to the change in the parameters, i.e. the performance for  $\sigma_x = \sigma_y = 0.1m$  is better than for  $\sigma_x = \sigma_y = 0.15m$ , which in turn is better than the performance for  $\sigma_x = \sigma_y = 0.2m$  and so on.

## **2.7 Conclusions and Future Work**

The extended target problem of tracking a single 2D convex-shaped target with known dynamics and without clutter was addressed. The problem of self-occlusion was also considered. A framework for tracking single extended targets was developed based on the existing point multitarget tracking framework. An algorithm was implemented under the current framework and tested for various scenarios. The shape consists only of the vertices describing the boundary and not the center (since the true center may be unobservable). The number of parameters describing the shape can be dynamically updated when more parts of the target become visible over time.

The performance of the algorithm was tested for various scenarios, including different measurement densities, measurement noise covariances, and shapes. The proposed algorithm was found to be robust under all these conditions.

The algorithm is limited by the convex polytope shape. Extension of the algorithm to 2D non-convex shapes and further to 3D shapes are to be attempted as part of future work. Since target rotation occurs frequently in practice, it becomes necessary to include target rotation as well in the dynamics for future work. Clutter can also be incorporated by using data association algorithms that support clutter. The parameters such as the measurement noise covariance and density of measurements as assumed to be known even though they can be unknown in reality. It is useful to include the parametric uncertainty in future work. As it is seen from Table (2.1), the runtime for the proposed algorithm is significantly greater than the existing Gaussian Process. Advanced association techniques such as PMHT (as opposed to the current nearest neighbour association) are to be studied to improve the runtime performance. Moreover, various data association, filtering ( depending on the target and measurement model), and track management techniques can be investigated under the proposed framework for single extended target tracking. Further, algorithms can be designed depending on the tradeoff in the performance based on the scenario.

## Bibliography

- [1] Johann Wolfgang Koch. Bayesian approach to extended object and cluster tracking using random matrices. *IEEE Transactions on Aerospace and Electronic Systems*, 44(3):1042–1059, 2008. doi: 10.1109/TAES.2008.4655362.
- [2] Marcus Baum and Uwe Hanebeck. Tracking an extended object modeled as an axis-aligned rectangle. pages 2422–2427, 01 2009.
- [3] Shishan Yang and Marcus Baum. Tracking the orientation and axes lengths of an elliptical extended object. *IEEE Transactions on Signal Processing*, 67(18):4720–4729, 2019. doi: 10.1109/TSP.2019.2929462.
- [4] Jian Lan and X. Rong Li. Tracking of maneuvering non-ellipsoidal extended object or target group using random matrix. *IEEE Transactions on Signal Processing*, 62(9):2450–2463, 2014. doi: 10.1109/TSP.2014.2309561.
- [5] Ben Liu, Ratnasingham Tharmarasa, Rahim Jassemi, Daly Brown, and Thia Kirubarajan. Extended target tracking with multipath detections, terrain-constrained motion model and clutter. *IEEE Transactions on Intelligent Transportation Systems*, 22(11):7056–7072, 2021. doi: 10.1109/TITS.2020.3001174.
- [6] Karl Granstrom, Marcus Baum, and Stephan Reuter. Extended object tracking: Introduction, overview and applications, 2017.

## BIBLIOGRAPHY

---

- [7] Niklas Wahlstrom and Emre Ozkan. Extended target tracking using gaussian processes. *IEEE Transactions on Signal Processing*, 63(16):4165–4178, 2015. doi: 10.1109/TSP.2015.2424194.
- [8] Marcus Baum and Uwe D. Hanebeck. Shape tracking of extended objects and group targets with star-convex rhms. In *14th International Conference on Information Fusion*, pages 1–8, 2011.
- [9] Xu Tang, Mingyan Li, Ratnasingham Tharmarasa, and Thiagalingam Kirubarajan. Seamless tracking of apparent point and extended targets using gaussian process pmht. *IEEE Transactions on Signal Processing*, 67(18):4825–4838, 2019. doi: 10.1109/TSP.2019.2932873.
- [10] Yunfei Guo, Yong Li, Ratnasingham Tharmarasa, Thiagalingam Kirubarajan, Murat Efe, and Bahadir Sarikaya. Gp - pda filter for extended target tracking with measurement origin uncertainty. *IEEE Transactions on Aerospace and Electronic Systems*, 55(4):1725–1742, 2019. doi: 10.1109/TAES.2018.2875555.
- [11] Antonio Zea, Florian Faion, Marcus Baum, and Uwe D. Hanebeck. Level-set random hypersurface models for tracking nonconvex extended objects. *IEEE Transactions on Aerospace and Electronic Systems*, 52(6):2990–3007, 2016. doi: 10.1109/TAES.2016.130704.
- [12] Abdullahi Daniyan, Sangarapillai Lambotharan, Anastasios Deligiannis, Yu Gong, and Wen-Hua Chen. Bayesian multiple extended target tracking using labeled random finite sets and splines. *IEEE Transactions on Signal Processing*, 66(22):6076–6091, 2018.

## BIBLIOGRAPHY

---

- [13] Hauke Kaulbersch, Jens Honer, and Marcus Baum. A cartesian b-spline vehicle model for extended object tracking. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 1–5, 2018. doi: 10.23919/ICIF.2018.8455717.
- [14] Art Owen. Empirical likelihood. 01 2002. doi: 10.1002/0471667196.ess0629.pub2.
- [15] Tao-Lin Lee, Rohan Attele, and Keh-Wei Chen. Maximum likelihood estimation in convex hull models. *Communications in Statistics - Theory and Methods*, 19(10):3759–3777, 1990. doi: 10.1080/03610929008830412. URL <https://doi.org/10.1080/03610929008830412>.
- [16] Y. Bar-Shalom, P.K. Willett, and X. Tian. *Tracking and Data Fusion: A Handbook of Algorithms*. YBS Publishing, 2011. ISBN 9780964831278. URL <https://books.google.ca/books?id=2a0iuAAACAAJ>.
- [17] Aytac Altan and Rifat Hacıoğlu. Model predictive control of three-axis gimbal system mounted on uav for real-time target tracking under external disturbances. *Mechanical Systems and Signal Processing*, 138:106548, 2020. ISSN 0888-3270. doi: <https://doi.org/10.1016/j.ymssp.2019.106548>. URL <https://www.sciencedirect.com/science/article/pii/S0888327019307691>.
- [18] Y. Bar-Shalom and X.R. Li. *Multitarget-multisensor Tracking: Principles and Techniques*. Yaakov Bar-Shalom, 1995. ISBN 9780964831209.
- [19] Yaakov Bar-Shalom, Fred Daum, and Jim Huang. The probabilistic data association filter. *IEEE Control Systems Magazine*, 29(6):82–100, 2009. doi: 10.1109/MCS.2009.934469.

## BIBLIOGRAPHY

---

- [20] S.S. Blackman. Multiple hypothesis tracking for multiple target tracking. *IEEE Aerospace and Electronic Systems Magazine*, 19(1):5–18, 2004. doi: 10.1109/MES.2004.1263228.
- [21] Yaakov Bar-Shalom, X.-Rong Li, and Thia Kirubarajan. *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*. 01 2004. ISBN 047141655X. doi: 10.1002/0471221279.ch11.
- [22] Simon J. Julier and Jeffrey K. Uhlmann. New extension of the Kalman filter to nonlinear systems. In Ivan Kadar, editor, *Signal Processing, Sensor Fusion, and Target Recognition VI*, volume 3068, pages 182 – 193. International Society for Optics and Photonics, SPIE, 1997. doi: 10.1117/12.280797. URL <https://doi.org/10.1117/12.280797>.
- [23] *Beyond the Kalman Filter: Particle Filters for Tracking Applications Particle Filters for Tracking Applications*. Artech House, 2004. ISBN 9781580536318.
- [24] *Convex Hulls*, pages 243–258. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-77974-2. doi: 10.1007/978-3-540-77974-2\_11. URL [https://doi.org/10.1007/978-3-540-77974-2\\_11](https://doi.org/10.1007/978-3-540-77974-2_11).
- [25] convhull matlab. <https://www.mathworks.com/help/matlab/ref/convhull.html>, .
- [26] convexhull python. <https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.ConvexHull.html>, .

## Chapter 3

# Extended Target Tracking using ET-PMHT for 3D Convex Polytope Shapes with Partial Visibility

The content of this chapter is ready to be submitted to a Journal as of December 8, 2023.

---

Mannari, Prabhanjan and Tharmarasa, Ratnasingham and Kirubarajan, Thiagalingam, "Extended Target Tracking using ET-PMHT for 3D Convex Polytope Shapes with Partial Visibility"

---

# **Extended Target Tracking using ET-PMHT for 3D Convex Polytope Shapes with Partial Visibility**

## **Abstract**

This article discusses the problem of tracking a single 3D extended target (or widely separated targets) with convex polytope shape when the target may only be partially visible. An extended target (as opposed to a point target) is one that may generate multiple measurements in a single frame. With the advent of high resolution sensors (such as LiDAR) the targets need to be considered as extended targets and their shape as well as kinematics need to be estimated. The extended target may only be partially visible (self-occlusion) and the measurements occur only from the visible parts of the target. In this work, different parts of a single extended target are assumed to be different targets constrained by the rigid body motion of the whole target and the multitarget tracking framework is used to handle the tracking. The target shape is described using a convex hull represented by its vertices and a Delaunay triangulation. The point target PMHT is modified to develop a single extended target PMHT (ET-PMHT) joint association and filtering by assuming that the face triangulations are separate targets. Face management is incorporated into the algorithm to delete erroneous faces and the algorithm is able to add new faces to refine the shape estimate. The framework is able to handle self-occlusion (partial visibility) by associating measurements only

to the visible parts of the target and vary the number of state parameters used to describe the shape. The performance of the algorithm is compared with 3D Gaussian process under various scenarios and RMSE of the center, velocity and IoU metrics are used to quantify the performance.

**Keywords:** *Extended Target, Self-Occlusion, Expectation Maximization, Probabilistic Multiple Hypothesis Tracker, 3D Convex Hull*

## **3.1 Introduction**

Traditional target tracking algorithms are designed assuming point targets, i.e. each target can generate at most one measurement per frame. The objective is to track the kinematics of possibly multiple targets. The problems of handling clutter, multitarget association and addition/deletion of targets are addressed in the traditional methods. However, with the advent of high resolution sensors (such as automotive radar, LiDAR), targets can produce multiple measurements across the sensor resolution cells in a single frame. Such targets are termed as extended targets and their shape as well as kinematics need to be estimated. This issue cannot be handled by traditional point target tracking algorithms and hence Extended Target (ET) algorithms have been developed for targets with finite extent (as opposed to zero extent of point targets).

The main challenge in extended target tracking is to estimate the state - shape and kinematics with the joint uncertainty in the state. The same set of measurements over time can be produced by targets with a combination of distinct shapes

and kinematics. One of the other issues that needs to be handled is the problem of visibility. Depending on the sensor-target geometry, only some parts of the target may be visible to the sensor and measurements are generated only from the visible parts. Self-occlusion occurs when some parts of the target are blocked from the sensor's view by the same target. When a target is blocked from the sensor's view by a different target, it is termed as mutual occlusion. Figure 3.1 shows an example of self-occlusion of convex shaped target from the view of the point  $S$ . Depending on the scenario, the entire target shape may not be observable due to occlusion even with multiple frames of measurements.

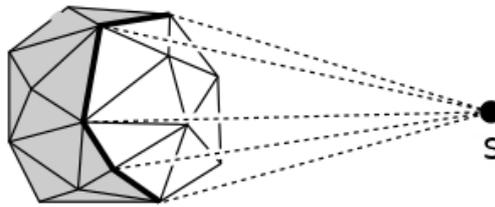


FIGURE 3.1: Example of self-occlusion. Visible faces are transparent and faces that are not visible are shaded. [1]

In [2],[3] shapes such as rectangles and ellipses are used for the extended target and the measurements over time are used to estimate the parameters such as the length of major/minor axes and the lengths of the sides of the rectangle. The Random Matrix approach is presented in [4] where the target shape is described by a symmetric positive semi-definite (PSD) matrix. More specifically, the measurements are generated around the center with a covariance given by the random matrix. The random matrix is assumed to be inverse Wishart distributed, which is the conjugate prior and an analytical Bayesian solution is derived. The algorithm is suited particularly for ellipsoidal shapes. [5] extends the Random Matrix

approach to non ellipsoidal shapes by using a combination of multiple ellipsoids to describe the shape. In [6], multipath detections and clutter issues are addressed in a terrain constrained environment using the Random Matrix model.

A radial basis function which follows a Gaussian Process is used in [7],[8] to describe a variety of star convex shapes for the extended target in 2D and 3D, respectively. The squared exponential function with appropriate distance function (2D/3D) is used to describe the covariance between different parts of the extended target. An analytical expression is derived for the Jacobian using an appropriate local measurement model and an EKF is developed for tracking extended targets. More details of [8] approach can be found in [4.6.1]. The same target can be viewed as a point target or an extended target depending on the sensor target geometry. [9] addresses this problem using the Gaussian process model and using a Poisson rate for the number of measurements. The Gaussian process model is extended to handle clutter in [10].

In [11] and [12], Random Hypersurface models are used to represent the target shape. Simple shapes such as rectangle, circle or ellipse can be used or more complicated shapes can be accommodated using a polygonal descriptor. A pseudo-measurement model is used with source location uncertainty to represent the measurements from the target. Both Random Hypersurface and Gaussian process models can handle measurements occurring from the interior of the target using a scaling factor. Splines are yet another model to represent the target shape addressed in [13],[14]. A comprehensive overview of the literature on extended target tracking is presented in [15].

In the current work, the different parts of the extended target are assumed to be different targets and the extended target tracking framework [16] developed based on the point multitarget tracking (MTT) framework is used to develop an algorithm to track 3D extended targets with convex polytope shape. The existing literature does not consider the visibility issues i.e., they assume measurements are generated from all parts of the target at all times, including those that are not visible to the sensor. In the present work, partially visible targets are handled by associating the measurements at certain time to only the visible parts of the target at that time. Due to self-occlusion, the center of the target may not be observable even over multiple frames of measurements. As such, the current work abandons the center of the target as a part of the state and only the boundary of the target is used to describe its shape, whereas in most existing algorithms, the center of the target is a part of the state to be estimated. Using the ET framework, the proposed algorithm is able to vary the number of parameters used to describe the shape as more parts of the target become visible over time.

The target shape is represented only by its boundary using a convex hull using its vertices and a Delaunay triangulation. The mesh triangulation to represent objects of different shapes is a common technique used in various fields of engineering such as structural analysis, fluid mechanics and electromagnetics [17]. Further, each point on the face triangulation can be represented uniquely as a linear combination of the vertices as they form a basis for 3D space. Hence, it is advantageous to represent the target shape in terms of face triangulations rather than polygonal faces. The standard Probabilistic Multiple Hypothesis Tracker (PMHT) is modified by assuming that the different face triangulations are separate targets

to develop Extended Target PMHT. Face management technique is used to delete faces with low probability and new faces can be added for measurements that do not fall inside the validation gate of the current estimate.

The main contributions of the paper are

1. The target shape model presented is able to describe a variety of 3D convex polytope shapes. The target shape is described by its vertices and the faces are described using a Delaunay triangulation.
2. A linear measurement model is developed for the target shape and the self-occlusion is incorporated.
3. Using the ET framework, an Extended Target PMHT (ET-PMHT) joint association and filtering is derived based on the point target PMHT.
4. An algorithm is presented which incorporates the ET-PMHT to track 3D extended targets with convex polytope shape while handling partial visibility.
5. The performance of the algorithm is compared with 3D Gaussian Process and various metrics are presented to indicate the performance.

The paper is organized as follows - Section II contains the problem description with the target and measurement models. Section III provides a review of the framework used for extended target tracking developed in [16]. Equations for ET-PMHT filtering are derived in Section IV and the complete algorithm is described in Section V. Section VI contains the simulation results and the conclusions are presented in Section VII.

## 3.2 Problem Description

The problem of tracking a single extended target (or widely separated targets) in 3D is discussed. The target is assumed to have a 3D convex polytope shape and is assumed to be a rigid body. The shape as well as the kinematics of the extended target are to be estimated using measurements over time. The visibility issues are considered as well - measurements only occur from the faces of the target visible to the sensor. The target is assumed to move according to a known dynamics model (nearly constant velocity model). The target state  $\mathbf{x}$  consists of a shape descriptor  $X_S$  which consists of the vertices of the target shape,  $X_{DT}$  which describes the edges/faces using a Delaunay Triangulation and a kinematics descriptor  $v$ . The matrix  $X_{DT}$  describing the edges using the vertices is not included directly in the state vector denoted by  $\mathbf{x}$ .

$$\mathbf{x} = \begin{bmatrix} X_S \\ v \end{bmatrix} \quad (3.1)$$

### 3.2.1 Target Model

The shape descriptor  $X_S$  for a target with  $N_X$  vertices can be written as  $\begin{bmatrix} p_1^T & p_2^T & p_3^T & \dots & p_{N_X}^T \end{bmatrix}^T$  where  $p_i$  are vertex points in 3D space, with  $p_i = \begin{bmatrix} p_i^x & p_i^y & p_i^z \end{bmatrix}^T$  denoting the  $x, y, z$  coordinates. The edges between the vertices are described in  $X_{DT}$  using a Delaunay triangulation of the convex polytope shape. The surface of the target is divided into non-overlapping triangles. Each row of  $X_{DT}$  consists

of the indices of the vertices of these triangles. Notice that the Delaunay triangulation of a 3D convex shape is not unique. The total target state is then given by

$$\begin{aligned}
 X_S &= \left[ p_1^x \ p_1^y \ p_1^z \ p_2^x \ \dots \ p_{N_X}^z \right]^T \\
 v &= \left[ v_x \ v_y \ v_z \right]^T \\
 \mathbf{x} &= \left[ p_1^T \ p_2^T \ \dots \ p_{N_X}^T \ v_x \ v_y \ v_z \right]^T
 \end{aligned} \tag{3.2}$$

The matrix  $X_{DT}$  for figure [3.2] is given by

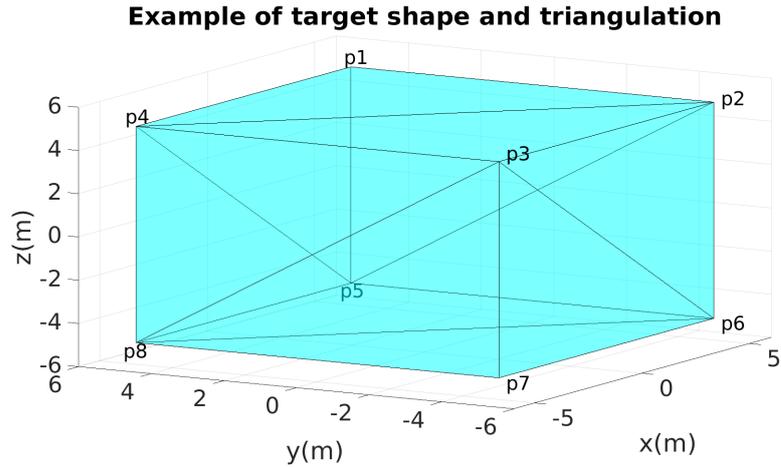


FIGURE 3.2: Example of a target as a cube and its Delaunay triangulation

$$X_{DT} = \begin{bmatrix} 1 & 2 & 5 \\ 1 & 4 & 2 \\ 1 & 5 & 4 \\ \vdots & \vdots & \vdots \\ 6 & 7 & 8 \end{bmatrix}$$

Let  $T$  be the sampling time between frames  $t$  and  $t + 1$ . For nearly constant velocity NCV model, the transition matrix  $F(t)$  and the process noise covariance  $Q(t)$  is given by

$$F(t) = \begin{bmatrix} I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & \dots & T \cdot I_{3 \times 3} \\ 0_{3 \times 3} & I_{3 \times 3} & 0_{3 \times 3} & \dots & T \cdot I_{3 \times 3} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & \dots & I_{3 \times 3} \end{bmatrix} \quad Q(t) = \begin{bmatrix} \frac{1}{3}T^3 & \frac{1}{2}T^2 & 0 & 0 & 0 & 0 \\ \frac{1}{2}T^2 & T & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3}T^3 & \frac{1}{2}T^2 & 0 & 0 \\ 0 & 0 & \frac{1}{2}T^2 & T & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3}T^3 & \frac{1}{2}T^2 \\ 0 & 0 & 0 & 0 & \frac{1}{2}T^2 & T \end{bmatrix}$$

$$\Gamma(t) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.3)$$

The dynamics model is then written as

$$\mathbf{x}(t + 1) = F(t)\mathbf{x}(t) + \Gamma(t)\nu(t) \quad \text{where } \nu(t) \sim \mathcal{N}(0, Q(t)) \quad (3.4)$$

The dynamics model described above differs from the NCV model for a point target. Since the extended target is a rigid body, all the points (including the vertices) have the same kinematics. This is achieved by using a  $6 \times 1$  process noise vector with process noise covariance  $Q(t)$  (one element for position in each dimension and one element for velocity in each dimension) and matrix  $\Gamma(t)$ .

### 3.2.2 Measurement Model

Measurements are obtained only from the faces of the target visible to the sensor. The measurements are assumed to occur only from the surface of the target corrupted by zero mean Gaussian noise with known covariance. For convex polytope shapes, a face of the target is visible to the sensor only if the centroid of the target and the sensor location are on the opposite sides of the face. This is the problem of self-occlusion and is particularly apparent in the 3D scenario.

The number of measurements from a face  $m$  is Poisson distributed, with the average number of measurements  $N_f^m$  proportional to the area of the face  $A_m$ .

$$N_f^m = \rho \times A_m \tag{3.5}$$

where  $\rho$  is the number of measurements per unit area.

Let the target  $\mathbf{x}$  have  $M$  number of faces (i.e. number of rows in the  $X_{DT}$  matrix) with each face represented by vertices  $(p_{m,1}, p_{m,2}, p_{m,3})$  where  $m = 1$  to  $M$ . Suppose  $n_{t,m}$  measurements occur from face  $m$ , with the total number of measurements at time frame  $t$  being  $n_t$ , then each measurement  $z_t^{m,r}$  from face  $m$  can be

written as

$$z_t^{m,r} = \lambda_{m,r,1}p_{m,1} + \lambda_{m,r,2}p_{m,2} + \lambda_{m,r,3}p_{m,3} + w \quad (3.6)$$

$$w \sim \mathcal{N}(0, R) \text{ and } \sum_{o=1}^3 \lambda_{m,r,o} = 1, \lambda_{m,r,o} \geq 0$$

where  $R$  is the measurement noise covariance.

The measurement  $z_t^{m,r}$  on the surface of face  $m$  is a linear combination of the vertices of the face, with the constraint on the coefficients  $\lambda_{m,r,o}$  to sum to one. The source locations on the surface of the target described by the  $\lambda$ s are unknown. The total set of measurements occurring from the face  $m$  can be written in matrix form as

$$\begin{bmatrix} z_t^{m,1} \\ z_t^{m,2} \\ z_t^{m,3} \\ \vdots \\ z_t^{m,n_{t,m}} \end{bmatrix} = \begin{bmatrix} \lambda_{m,1,1}I_{3 \times 3} & \lambda_{m,1,2}I_{3 \times 3} & \lambda_{m,1,3}I_{3 \times 3} \\ \lambda_{m,2,1}I_{3 \times 3} & \lambda_{m,2,2}I_{3 \times 3} & \lambda_{m,2,3}I_{3 \times 3} \\ \lambda_{m,3,1}I_{3 \times 3} & \lambda_{m,3,2}I_{3 \times 3} & \lambda_{m,3,3}I_{3 \times 3} \\ \vdots & \vdots & \vdots \\ \lambda_{m,n_{t,m},1}I_{3 \times 3} & \lambda_{m,n_{t,m},2}I_{3 \times 3} & \lambda_{m,n_{t,m},3}I_{3 \times 3} \end{bmatrix} \begin{bmatrix} p_{m,1} \\ p_{m,2} \\ p_{m,3} \end{bmatrix} + w_m \quad (3.7)$$

The total measurement set from face  $m$  can be compactly written as-

$$\{Z_t^m = z_t^{m,r}\}_{r=1}^{n_{t,m}} = H_t^m \begin{bmatrix} p_{i1} \\ p_{i2} \\ p_{i3} \end{bmatrix} + w_m \quad (3.8)$$

$w_m$  is the stacked measurement noise vector for the face and is distributed as  $w_m \sim \mathcal{N}(0, R_m)$ , where  $R_m$  is the block diagonal matrix of  $R$  with the number of

entries corresponding to the number of measurements being addressed, i.e.  $n_{t,m}$ .

The total measurement set from all faces  $\{Z_t = z_t^{m,r}\}_{m=1}^M, r=1}^{n_t}$  can be written using the above notation as

$$Z_t = \mathcal{H}_t \mathbf{x} + w_t \quad (3.9)$$

$w_t$  is the stacked measurement noise vector for the measurements at time frame  $t$  and is distributed as  $w_t \sim \mathcal{N}(0, R_t)$ . In a similar manner as before,  $R_t$  is the block diagonal matrix of  $R$  with the number of entries being equal to the total number of measurements at the time, i.e.  $n_t$ . The rows of  $\mathcal{H}_t$  corresponding to certain measurement  $z_t^{m,r}$  from face  $m$  consists of the coefficients for the vertices of face  $m$  and the rest of the elements are zero, which are the vertices that do not correspond to the face and hence do not contribute to the measurement. Figure 3.3 shows an example of measurements occurring from the visible faces of the target.

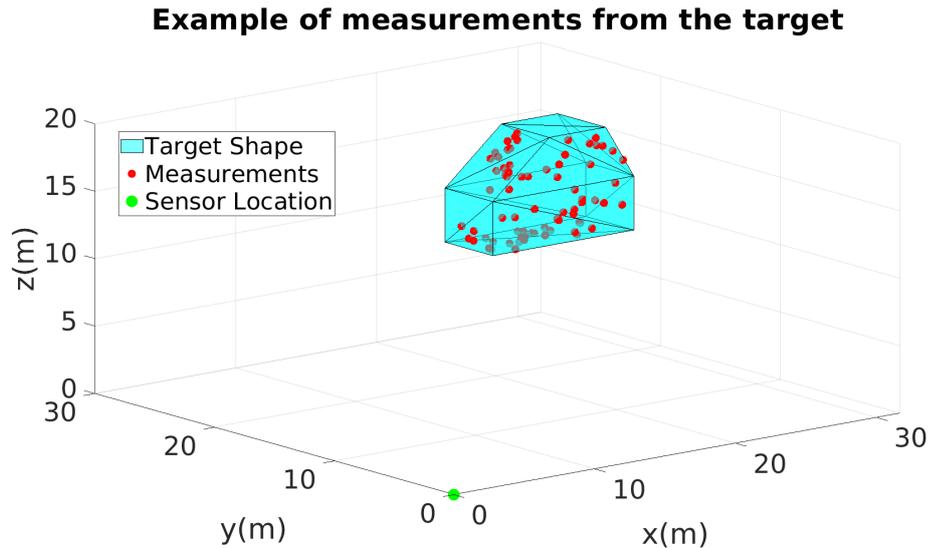


FIGURE 3.3: Example of measurements from visible faces of the target shape

### **3.3 MTT Framework for single extended target tracking**

The multitarget tracking framework has been extensively developed for point target tracking and various techniques have been introduced to handle different tracking scenarios. By treating different parts of a single extended target as multiple targets, the point MTT framework can be modified accordingly to tackle the extended target tracking problem [16]. The typical MTT framework consists of the following steps

- Gating - The measurements are validated using the current target estimates and grouped by the gates of the targets they fall into (single measurement can fall into multiple overlapping gates). The measurements not falling inside any gate are grouped separately and used for initialization.
- Association - Multitarget association techniques are used to associate the validated measurements to the targets or clutter.
- Filtering - The target estimates are filtered using the associated measurements.
- Track Initialization - The measurements which do not fall inside the gate of any current target are used to initialize new targets.
- Track Management - The track quality scores are calculated for the available tracks. Tracks can be termed as tentative, confirmed or deleted depending

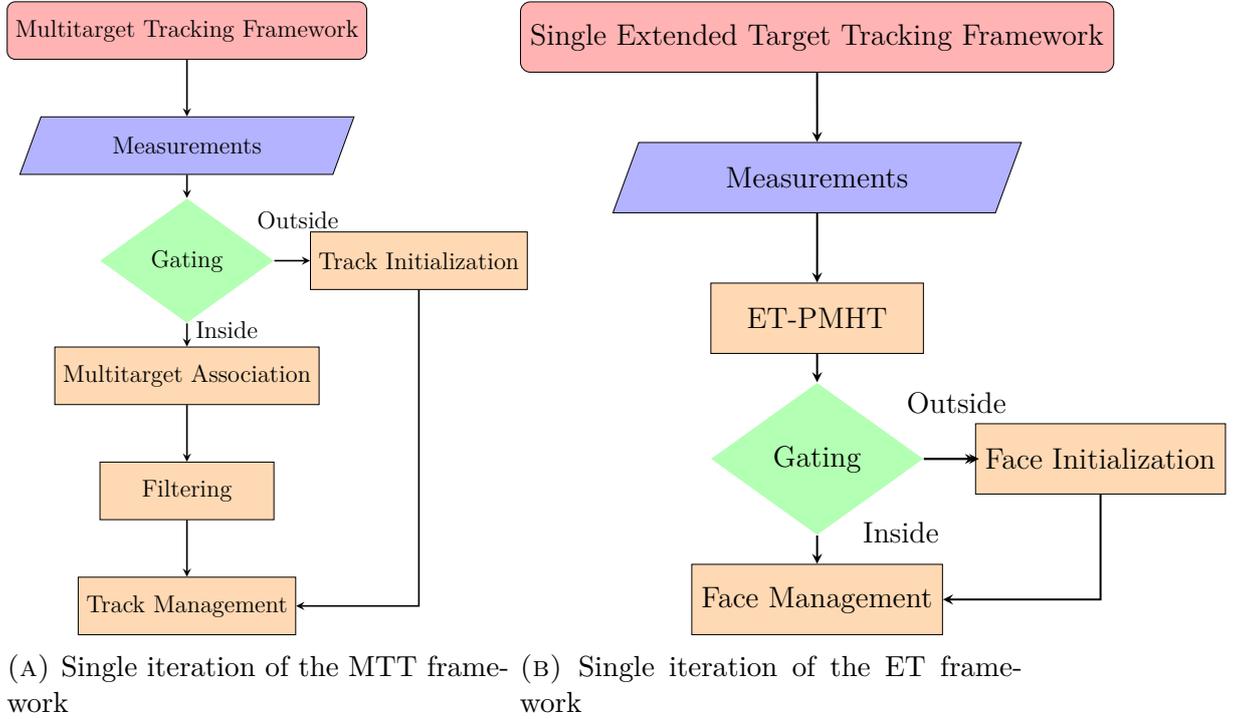


FIGURE 3.4: Flowchart of the frameworks

on the current track quality.

The general modification of the MTT framework for ET tracking has been described in [16] and the flowchart is given in Figure 3.4

## 3.4 Solution

### 3.4.1 Determination of the measurement function

Given the measurements  $Z = \{z^1, z^2, \dots, z^n\}$  and the face triangulation described by the points  $\{p_1, p_2, p_3\}$ , the measurement function relating the measurements to

the face is to be determined. The least squares method with linear constraints is used to estimate the measurement function. It is assumed that the measurement function for a certain measurement-face pair is independent of the other faces and measurements. The measurement equation for the  $r^{th}$  measurement from the face is given as

$$z^r = \lambda_{r,1}p_1 + \lambda_{r,2}p_2 + \lambda_{r,3}p_3 + w \quad \text{where } w \sim \mathcal{N}(0, R)$$

$$z^r = \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix} \begin{bmatrix} \lambda_{r,1} \\ \lambda_{r,2} \\ \lambda_{r,3} \end{bmatrix} + w = P\Lambda_r + w \quad (3.10)$$

Using the form of the measurement equation above, the estimation is cast as a constrained optimization problem

$$\hat{\Lambda}_r = \arg \min_{\Lambda_r} (z^r - P\Lambda_r)^T R^{-1} (z^r - P\Lambda_r) \quad (3.11)$$

subject to  $\lambda_{r,o} \geq 0$  and  $\sum_{o=1}^3 \lambda_{r,o} = 1$

The total measurement matrix can then be obtained in a manner similar to equation (3.7). It is to be noted that solution of the optimization is the coefficients for the point on the face triangulation closest to the measurement in terms of Equation (3.11) for the distance measure. Geometric methods can be used instead of expensive optimization routines to obtain the solution.

### 3.4.2 Expectation Maximization for static shape estimation: ET-PMHT initialization

Similar to the Expectation Maximization (EM) procedure applied to a static Gaussian mixture in [18], the EM method is used to obtain shape estimate from a set of measurements with unknown measurement to face association. This technique is used in shape initialization, i.e. no prior. The measurement set at frame 0 (initial time frame) is  $Z_0 = \{z_0^1, z_0^2, \dots, z_0^{n_0}\}$  and the unknown measurement to face association, the missing data is denoted by  $K_0 = \{k_0^1, k_0^2, \dots, k_0^{n_0}\}$ . Each element  $k_t^r$  denotes that the  $r^{\text{th}}$  measurement at time frame  $t$  is associated to the  $k_t^{r^{\text{th}}}$  face and hence  $k_t^r$  takes values from 1 to  $M$  where  $M$  is the number of faces in the estimate. The initial estimate provided to the procedure is denoted by  $X_S(0) = \left[ (p_0^1)^T \quad (p_0^2)^T \quad \dots \quad (p_0^{N_0})^T \right]^T$  with the face triangulations  $F_0 = \{f_0^1, f_0^2, \dots, f_0^M\}$ . The face triangulations are encoded by  $X_{DT}(0)$  of size  $M \times 3$ . Each element of the face triangulation  $f_t^m$  is  $9 \times 1$  vector consisting of the three 3D points that are vertices of the triangle being described. Assuming that the measurements occur independently of one another, the complete data likelihood can be written as

$$p_{comp}(Z_0, K_0, X_S(0)) = \prod_{r=1}^{n_0} \mathcal{N}(z_0^r; h_0^{k_0^r, r}(f_0^{k_0^r}), R) \quad (3.12)$$

$h^{k_t^r, r}$  is the measurement function for the  $r^{\text{th}}$  measurement at time frame  $t$  occurring from the face  $f_t^{k_t^r}$ . Since the measurement function is calculated separately for each face-measurement pair, it is independent of the association and is hence not included in the missing data. It is however, a function of the shape state  $X_S(0)$

but is assumed to be a constant estimate with each iteration. The measurement function for the current iteration is estimated using the estimated state  $\hat{X}_S(0)$  from the previous iteration.

The conditional probability of the missing data is calculated in a similar manner as in [18]

$$p_{miss}(K_0|Z_0, \hat{X}_S(0)) = \prod_{r=1}^{n_0} \frac{\mathcal{N}(z_0^r; h^{k_0^r, r}(f_0^{k_0^r}), R)}{\sum_{m=1}^M \mathcal{N}(z_0^r; h^{m, r}(f_0^m), R)} = \prod_{r=1}^{n_0} p(k_0^r|z_0^r, \hat{X}_S(0)) \quad (3.13)$$

Denoting the weights for the  $m^{th}$  face and  $r^{th}$  measurement pair as  $w_0^{m, r}$ ,

$$w_0^{m, r} = p(k_0^r = m|z_0^r, \hat{X}_S(0)) \quad (3.14)$$

The auxiliary function is the expected value of the logarithm of the complete data likelihood over the missing data. It can be calculated as

$$\begin{aligned} \mathcal{L}(X_S(0)|\hat{X}_S(0)) &= E_{K_0} \left\{ \log(p_{comp}(Z_0, K_0, X_S(0))) \right\} \\ &= -\frac{1}{2} \sum_{r=1}^{n_0} \log |2\pi R| - \frac{1}{2} \sum_{m=1}^M \sum_{r=1}^{n_0} \underbrace{(z_0^r - h_0^{m, r}(f_0^m))^T R^{-1} (z_0^r - h_0^{m, r}(f_0^m)) w_0^{m, r}} \end{aligned} \quad (3.15)$$

The first term is independent of the state to be estimated and is dropped from the auxiliary function in the further steps. The measurement function is linear (from equation (3.7)), i.e.  $h_0^{m, r}(f_0^m) = h_0^{m, r} f_0^m$ . The underlined part of the equation can be simplified as follows

$$\begin{aligned} \sum_{r=1}^{n_0} (z_0^r - h_0^{m, r}(f_0^m))^T R^{-1} (z_0^r - h_0^{m, r}(f_0^m)) w_0^{m, r} \\ = (Z_0 - H_0^m f_0^m)^T (R_0^m)^{-1} (Z_0 - H_0^m f_0^m) \end{aligned} \quad (3.16)$$

where

$$Z_0 = \begin{bmatrix} z_0^1 \\ z_0^2 \\ \vdots \\ z_0^{n_0} \end{bmatrix}, \quad H_0^m = \begin{bmatrix} h_0^{m,1} \\ h_0^{m,2} \\ \vdots \\ h_0^{m,n_0} \end{bmatrix}, \quad (R_0^m)^{-1} = \begin{bmatrix} w_0^{m,1} R^{-1} & & & & \\ & w_0^{m,2} R^{-1} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & w_0^{m,n_0} R^{-1} \end{bmatrix} \quad (3.17)$$

Face  $f_0^m$  can be written in terms of the state  $X_S(0)$  as  $f_0^m = F^{mx} X_S(0)$ . Let  $N_X(t)$  be the number of vertices in the shape estimate at time frame  $t$ . Matrix  $F^{mx}$  is of size  $9 \times 3N_X(t)$ , which extracts the vertices describing the face from the points in the state.  $F^{mx}$  is zeros everywhere except the sets of columns (each set is 3 consecutive columns) given by the  $m^{\text{th}}$  row of  $X_{DT}$ . For each set of rows  $s$  (each set is 3 consecutive rows) the  $X_{DT}(m, s)$  set of columns contains  $I_{3 \times 3}$ . Suppose the face  $f^m$  consists of points  $\begin{bmatrix} p_1^T & p_3^T & p_5^T \end{bmatrix}^T$  from  $X_S = \begin{bmatrix} p_1^T & p_2^T & p_3^T & p_4^T & p_5^T \end{bmatrix}^T$  then  $F^{mx}$  is given by-

$$f^m = \begin{bmatrix} I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & I_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & 0_{3 \times 3} & I_{3 \times 3} \end{bmatrix} X_S \quad (3.18)$$

The auxiliary function, dropping terms independent of the state, can then be written directly in terms of the state as

$$\mathcal{L}(X_S(0) | \hat{X}_S(0)) = -\frac{1}{2} \sum_{m=1}^M (Z_0 - H_0^m F^{mx} X_S(0))^T (R_0^m)^{-1} (Z_0 - H_0^m F^{mx} X_S(0)) \quad (3.19)$$

Since the auxiliary function is a sum of (at least) negative semi-definite matrices, the maximum can be found by taking the derivative of the auxiliary function with respect to the state and equating it to zero

$$\begin{aligned}
 \nabla_{X_S(0)} \mathcal{L}(X_S(0) | \hat{X}_S(0)) &= 0 \\
 \implies \sum_{m=1}^M (H_0^m F^{mx})^T (R_0^m)^{-1} (Z_0 - H_0^m F^{mx} X_S(0)) &= 0 \\
 \implies \left( \sum_{m=1}^M (H_0^m F^{mx})^T (R_0^m)^{-1} H_0^m F^{mx} \right) X_S(0) &= \left( \sum_{m=1}^M (H_0^m F^{mx})^T (R_0^m)^{-1} \right) Z_0
 \end{aligned} \tag{3.20}$$

The EM procedure is followed until the difference in the values of the auxiliary function between iterations  $l$  and  $l + 1$  falls below a certain threshold

$$\frac{1}{n_0} \left| \mathcal{L}(X_S(0) | \hat{X}_S(0))|_{l+1} - \mathcal{L}(X_S(0) | \hat{X}_S(0))|_l \right| < \tau_I \tag{3.21}$$

### 3.4.3 Expectation Maximization for dynamic state estimation: ET-PMHT update

A procedure similar to the above procedure is developed for dynamic state estimation, where the state evolves according to known dynamics. The data for the Expectation Maximization consists of the total measurement set for time frames  $t = 1$  to  $T$ :  $\mathbb{Z} = \{Z_1, Z_2, \dots, Z_T\}$ , the states to be estimated:  $\mathbb{X} = \{\mathbf{x}(0), \mathbf{x}(1), \dots, \mathbf{x}(T)\}$  and the missing data of the association terms is denoted by:  $\mathbb{K} = \{K_1, K_2, \dots, K_T\}$ . The individual terms of the total measurement set

$\mathbb{Z}$  are the measurement sets for each time  $t$ :  $Z_t = \{z_t^1, z_t^2, \dots, z_t^{n_t}\}$  and the individual terms of the  $\mathbb{X}$  are the states each time.  $z_t^r$  is the  $r^{\text{th}}$  measurement at time  $t$  of size  $3 \times 1$  consisting of  $(x, y, z)$  co-ordinates.  $f_t^m$  is the state of the  $m^{\text{th}}$  face at time  $t$  of size  $9 \times 1$  consisting of the three vertices describing the face triangulation  $(p_1(t), p_2(t), p_3(t))$  each of size  $3 \times 1$ . The individual terms of  $\mathbb{K}$  are  $K_t = \{k_t^1, k_t^2, \dots, k_t^{n_t}\}$  in which  $k_t^r$  denotes that the  $z_t^r$  is associated to the face  $f_t^{k_t^r}$ . Hence  $k_t^r$  takes values in the set  $\{1, 2, \dots, M\}$ . It must be noted that the number of faces or the face triangulations do not change during the procedure. The complete data likelihood is then

$$p_{comp}(\mathbb{X}, \mathbb{Z}, \mathbb{K}) = p(\mathbb{Z}|\mathbb{X}, \mathbb{K})p(\mathbb{K}|\mathbb{X})p(\mathbb{X}) \quad (3.22)$$

where,  $p(\mathbb{Z}|\mathbb{X}, \mathbb{K})$  consists of the measurement likelihood terms -

$$p(\mathbb{Z}|\mathbb{X}, \mathbb{K}) = \prod_{t=1}^T \prod_{r=1}^{n_t} \mathcal{N}(z_t^r; h_t^{k_t^r, r}(f_t^{k_t^r}), R) \quad (3.23)$$

$p(\mathbb{K}|\mathbb{X})$  denotes the probability of the association of the number of measurements to the faces of the given state estimates. This term is handled in the face management step of update (3.5.2) and only the faces with  $p(\mathbb{K}|\mathbb{X})$  greater than a threshold are used for the EM procedure.

$p(\mathbb{X})$  describes the evolution of the states according to the known dynamics independent of the measurements. Let  $x_t^m = \begin{bmatrix} f_t^m \\ v(t) \end{bmatrix}$ .  $p(\mathbb{X})$  can then be written as -

$$p(\mathbb{X}) = \prod_{m=1}^M p(x_0^m) \prod_{t=1}^T [p(x_t^m | x_{t-1}^m)] \quad (3.24)$$

The conditional probability of the missing data is given by

$$p(\mathbb{K}|\mathbb{Z}, \hat{\mathbb{X}}) = \prod_{t=1}^T \prod_{r=1}^{n_t} p(k_t^r | z_t^r, \hat{\mathbf{x}}(t)) \quad (3.25)$$

The conditional probability of associating  $r^{th}$  measurement to the  $m^{th}$  face is denoted as  $w_t^{m,r} = p(k_t^r = m | z_t^r, \mathbf{x}(t))$ .

The auxiliary function is then -

$$\begin{aligned} \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}}) &= E_{\mathbb{K}} \left( \log(p_{comp}(\mathbb{X}, \mathbb{K}, \mathbb{Z})) \right) \\ &= E_{\mathbb{K}} \left( \sum_{t=1}^T \sum_{r=1}^{n_t} \log\{\mathcal{N}(z_t^r; h_t^{k_t^r, r}(f_t^{k_t^r}), R)\} + \sum_{m=1}^M \left( \log\{p(x_0^m)\} + \sum_{t=1}^T \log\{p(x_t^m | x_{t-1}^m)\} \right) \right) \end{aligned} \quad (3.26)$$

The state evolution terms are independent of the association. The expected value of the measurement likelihood terms is calculated in a similar way as before. Note that the determinant terms in the measurement likelihood have been dropped as they are independent of the states to be estimated.

$$\begin{aligned} \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}}) &= -\frac{1}{2} \sum_{m=1}^M \sum_{t=1}^T \sum_{r=1}^{n_t} \left( (z_t^r - h_t^{m,r}(f_t^m))^T R^{-1} (z_t^r - h_t^{m,r}(f_t^m)) w_t^{m,r} \right) \\ &+ \sum_{m=1}^M \left( \log\{p(x_0^m)\} + \sum_{t=1}^T \log\{p(x_t^m | x_{t-1}^m)\} \right) \end{aligned} \quad (3.27)$$

The underlined part of the above equation can be simplified in a similar way as done previously. The measurement function is assumed to be linear  $h_t^{m,r}(f_t^m) = h_t^{m,r} f_t^m$ .

$$\begin{aligned} \sum_{r=1}^{n_t} \left( (z_t^r - h_t^{m,r}(f_t^m))^T R^{-1} (z_t^r - h_t^{m,r}(f_t^m)) w_t^{m,r} \right) \\ = (Z_t - H_t^m f_t^m)^T (R_t^m)^{-1} (Z_t - H_t^m f_t^m) \end{aligned} \quad (3.28)$$

where

$$H_t^m = \begin{bmatrix} h_t^{m,1} \\ h_t^{m,2} \\ \vdots \\ h_t^{m,n_t} \end{bmatrix}, \quad Z_t = \begin{bmatrix} z_t^1 \\ z_t^2 \\ \vdots \\ z_t^{n_t} \end{bmatrix}, \quad (R_t^m)^{-1} = \begin{bmatrix} w_t^{m,1} R^{-1} & & & \\ & w_t^{m,2} R^{-1} & & \\ & & \ddots & \\ & & & w_t^{m,n_t} R^{-1} \end{bmatrix} \quad (3.29)$$

Each face  $f_t^m$  can be written in terms of the state  $\mathbf{x}(t)$  as  $f_t^m = F^{mx} \mathbf{x}(t)$ . Note the overloading of the notation for  $F^{mx}$ . When used with the state instead of the shape  $X_S(t)$ ,  $F^{mx}$  has 3 extra zero columns. Similarly we can define  $F^{mxv}$  such that  $x_t^m = F^{mxv} \mathbf{x}(t)$ . Assuming the prior  $p(\mathbf{x}(0))$  to be Gaussian  $\mathcal{N}(\mathbf{x}(0); \bar{\mathbf{x}}(0), P(0))$ , and Gaussian dynamics according to Equation (3.4), i.e.  $p(\mathbf{x}(t)|\mathbf{x}(t-1)) = \mathcal{N}(\mathbf{x}(t); F(t)\mathbf{x}(t-1), Q(t))$ , the auxiliary function can be written as -

$$\begin{aligned} \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}}) = & -\frac{1}{2} \sum_{m=1}^M \sum_{t=1}^T \left( (Z_t - H_t^m F^{mx} \mathbf{x}(t))^T (R_t^m)^{-1} (Z_t - H_t^m F^{mx} \mathbf{x}(t)) \right) \\ & -\frac{1}{2} \sum_{m=1}^M \left( (\mathbf{x}(0) - \bar{\mathbf{x}}(0))^T (P_0^m)^{-1} (\mathbf{x}(0) - \bar{\mathbf{x}}(0)) + \right. \\ & \left. \sum_{t=1}^T (\mathbf{x}(t) - F(t)\mathbf{x}(t-1))^T (Q_t^m)^{-1} (\mathbf{x}(t) - F(t)\mathbf{x}(t-1)) \right) \end{aligned} \quad (3.30)$$

where  $(P_0^m)^{-1} = (F^{mxv})^T P(0)^{-1} F^{mxv}$  and  $(Q_t^m)^{-1} = (F^{mxv})^T Q(t)^{-1} F^{mxv}$ . Note that the time indices for  $F(t)$  and  $Q(t)$  have been dropped in the further steps since they are constant. To account for visibility issues in the filtering, the measurements are associated only to the faces visible at the time. Let  $M_t^{vis}$  be the set of indices of the faces visible at time  $t$ . The auxiliary function is modified accordingly as -

$$\begin{aligned} \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}}) = & -\frac{1}{2} \sum_{t=1}^T \sum_{m \in M_t^{vis}} \left( (Z_t - H_t^m F^{mx} \mathbf{x}(t))^T (R_t^m)^{-1} (Z_t - H_t^m F^{mx} \mathbf{x}(t)) \right) \\ & - \frac{1}{2} \sum_{m=1}^M \left( (\mathbf{x}(0) - \bar{\mathbf{x}}(0))^T (P_0^m)^{-1} (\mathbf{x}(0) - \bar{\mathbf{x}}(0)) \right. \\ & \left. + \sum_{t=1}^T (\mathbf{x}(t) - F\mathbf{x}(t-1))^T (Q_t^m)^{-1} (\mathbf{x}(t) - F\mathbf{x}(t-1)) \right) \quad (3.31) \end{aligned}$$

The auxiliary function is a sum of at least negative semi-definite quadratic forms and hence, the maximum is found by equating the first derivative to zero. The calculation of the derivatives is given in Appendix 3.8. The total linear system can be written as -

$$\begin{aligned} \begin{bmatrix} \sum_{m=1}^M (P_0^m)^{-1} + F^T (Q_0^m)^{-1} F & - \sum_{m=1}^M F^T (Q_0^m)^{-1} & 0 & 0 & \dots & 0 \\ - \sum_{m=1}^M (Q_0^m)^{-1} F & A_1 & - \sum_{m=1}^M F^T (Q_1^m)^{-1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & - \sum_{m=1}^m (Q_{T-1}^m)^{-1} F & A_T \end{bmatrix} \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{x}(1) \\ \vdots \\ \mathbf{x}(T) \end{bmatrix} \\ = \begin{bmatrix} \sum_{m=1}^M (P_0^m)^{-1} \bar{\mathbf{x}}(0) \\ B_1 Z_1 \\ \vdots \\ B_T Z_T \end{bmatrix} \quad (3.32) \end{aligned}$$

where

$$A_t = \sum_{m \in M_t^{vis}} (H_t^m F^{mx})^T (R_t^m)^{-1} (H_t^m F^{mx}) + \sum_{m=1}^M (Q_{t-1}^m)^{-1} + F^T (Q_t^m)^{-1} F \quad (3.33)$$

$$B_t = \sum_{m \in M_t^{vis}} (H_t^m F^{mx})^T (R_t^m)^{-1} \quad (3.34)$$

Using  $t - 1$  and  $t$  as the initial and final time frames, the equations for a single frame update can be obtained as follows -

$$\begin{aligned} & \begin{bmatrix} \sum_{m=1}^M (P_{t-1}^m)^{-1} + F^T (Q_{t-1}^m)^{-1} F & - \sum_{m=1}^M F^T (Q_{t-1}^m)^{-1} \\ - \sum_{m=1}^M (Q_{t-1}^m)^{-1} F & \sum_{m \in M_t^{vis}} (H_t^m F^{mx})^T (R_t^m)^{-1} (H_t^m F^{mx}) + \sum_{m=1}^M (Q_{t-1}^m)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}(t-1) \\ \mathbf{x}(t) \end{bmatrix} \\ & = \begin{bmatrix} \sum_{m=1}^M (P_{t-1}^m)^{-1} \mathbf{x}(t-1) \\ \sum_{m \in M_t^{vis}} (H_t^m F^{mx})^T (R_t^m)^{-1} Z_t \end{bmatrix} \quad (3.35) \end{aligned}$$

The terms of the auxiliary function that depend on  $\mathbf{x}(t)$  have the form of the Kalman filter negative log-likelihood equations. Hence, the Kalman Filter equations can be used for finding the updated covariance.

$$\begin{aligned} \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}})|_{\mathbf{x}(t)terms} &= -\frac{1}{2} \sum_{m \in M_t^{vis}} (Z_t - H_t^m F^{mx} \mathbf{x}(t))^T (R_t^m)^{-1} (Z_t - H_t^m F^{mx} \mathbf{x}(t)) \\ & - \sum_{m=1}^M (\mathbf{x}(t) - F \mathbf{x}(t-1))^T (Q_{t-1}^m)^{-1} (\mathbf{x}(t) - F \mathbf{x}(t-1)) \quad (3.36) \end{aligned}$$

The equations for the covariance update are then

$$P(t|t-1) = FP(t-1|t-1)F^T + \Gamma Q(t)\Gamma^T \quad (3.37)$$

$$S(t) = H(t)P(t|t-1)H(t)^T + R(t) \quad (3.38)$$

$$W(t) = P(t|t-1)H(t)^T S(t)^{-1} \quad (3.39)$$

$$P(t|t) = P(t|t-1) - W(t)S(t)^{-1}W(t)^T \quad (3.40)$$

where the total measurement matrix  $H(t)$  consists of vertically stacked  $H_t^m F^{mx}$  for visible  $m$  and  $R(t)$  is the block diagonal matrix with the blocks as  $R_t^m$  for visible  $m$ . The nonlinearity of the estimation is reflected in the updated covariance being a function of  $H(t)$  and  $R(t)$ , which in turn depend on the current estimate at time  $t$ .

The EM procedure is followed until the difference in the values of the auxiliary function between iterations  $l$  and  $l+1$  falls below a certain threshold  $\tau_U$ .

$$\frac{1}{n_t} \left| \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}})|_{l+1} - \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}})|_l \right| < \tau_U \quad (3.41)$$

In a similar manner, the equations to update only the velocity can be derived, while keeping the shape fixed.

$$\begin{aligned}
 & \begin{bmatrix} P_{vv}^{-1}(t-1) + Q_{vv}^{-1} & & -Q_{vv}^{-1} \\ & -Q_{vv}^{-1} & \\ & & Q_{vv}^{-1} + \sum_{m \in M_t^{vis}} \Gamma_{vv}^T (H_t^m F^{mx})^T (R_t^m)^{-1} (H_t^m F^{mx}) \Gamma_{vv} \end{bmatrix} \begin{bmatrix} v(t-1) \\ v(t) \end{bmatrix} \\
 &= \begin{bmatrix} P_{vv}^{-1}(t-1) \bar{v}(t-1) \\ \sum_{m \in M_t^{vis}} b_m \end{bmatrix} \tag{3.42}
 \end{aligned}$$

$$b_m = \Gamma_{vv}^T (H_t^m F^{mx})^T (R_t^m)^{-1} Z_t - \Gamma_{vv}^T (H_t^m F^{mx})^T (R_t^m)^{-1} (H_t^m F^{mx}) \mathbf{x}(t-1)$$

where  $P_{vv}, Q_{vv}$  are the blocks of the estimate covariance and process noise covariance for the velocity part of the state and  $\Gamma_{vv}$  is such that  $F\mathbf{x}(t) = \mathbf{x}(t) + \Gamma_{vv}v(t)$ .

### 3.4.4 Simplification of the convex hull

The basic convex hull operation does not take into account the uncertainty in the input points and this may result in a large number of faces in the output. To account for this uncertainty, a convex hull simplification operation is proposed. The smaller convex hull is calculated by removing points one by one from the original convex hull. If the point which is removed is within the validation gate of the smaller convex hull, then the smaller convex hull is considered as the original. It is to be noted that this is a greedy algorithm since the original convex hull is modified point by point as opposed to calculating the effect of all points and then modifying the original hull. The parameter used to define the validation region for the simplification process is denoted by  $\Gamma_S$ .

## 3.5 Algorithm

### 3.5.1 Initialization

The objective is to find an initial estimate of the state  $\hat{\mathbf{x}}(0) = \left[ \hat{X}_S(0)^T \quad \hat{v}_0^T \right]^T$  and  $\hat{X}_{DT}(0)$  given the measurements at time 0,  $Z_0 = \{z_0^1, z_0^2, \dots, z_0^{n_0}\}$ . The velocity is initialized with a given  $\hat{v}(0)$  and covariance  $P_v(0)$  independent of the shape. The below procedure is followed for estimating the initial shape until the cost (3.21) between successive iterations falls below a given threshold or a maximum number of iterations is reached. The initial estimate for the iterations is the visible faces of the convex hull of the measurements with pre-defined covariance  $P_S(0)$ .

1. Joint association and filtering - The EM procedure for static state estimation is termed as ET-PMHT for initialization (3.20) and is used for joint association and filtering. The ET-PMHT procedure does not change the face triangulations or the number of faces, but it may not maintain the convexity of the shape. The procedure, as such, can also be used with non-convex shapes or even disjoint shapes with appropriate triangulations.
2. Convex hull constraint - The estimate after ET-PMHT procedure may not be convex, and hence, the convex hull operation is used on the estimate to constrain the shape to be convex. The convex hull is simplified to reduce the number of faces.
3. Extracting the visible faces - The visible faces are extracted from the convex shape estimate and used as the initial estimate for the next iteration

The initialization pseudocode is given below in Algorithm 5

---

**Algorithm 3** Shape Initialization

---

```

procedure INITIALIZE_SHAPE( $Z(0)$ , iter_max)
     $X_S^0, X_{DT}^0 \leftarrow \text{get\_visible\_faces}(\text{simplified\_convex\_hull}(Z(0)))$ 
     $L_X, L_P, L_{DT}, L_C \leftarrow [], [], [], []$ 
    for  $l=1, l < \text{iter\_max}$  do
         $X_S^l, P_S^l, X_{DT}^l, C^l \leftarrow \text{ET-PMHT-initialization}(X_S^{l-1}, X_{DT}^{l-1}, Z(0))$ 
         $X_S^l, P_S^l, X_{DT}^l \leftarrow \text{simplified\_convex\_hull}(X_S^l, P_S^l, X_{DT}^l)$ 
         $L_X, L_P, L_{DT}, L_C \leftarrow [L_X \ X_S^l], [L_P \ P_S^l], [L_{DT} \ X_{DT}^l], [L_C \ C^l]$ 
         $X_S^l, X_{DT}^l \leftarrow \text{get\_visible\_faces}(X_S^l, X_{DT}^l)$ 
     $\hat{X}_S(0), P_S(0), \hat{X}_{DT}(0) \leftarrow \text{min\_cost\_estimate}(L_X, L_P, L_{DT}, L_C)$ 

```

---

### 3.5.2 Update

Given the inputs  $\hat{\mathbf{x}}(t), P(t), \hat{X}_{DT}(t)$  from time  $t$  and the measurement set  $Z(t+1)$  at time  $t+1$  the objective is to find the updated state estimate and its covariance -  $\hat{\mathbf{x}}(t+1), P(t+1), \hat{X}_{DT}(t+1)$ . The following steps are performed till the cost (3.41) between two consecutive state estimates falls below a given threshold or until a maximum number of iterations is reached.

1. Initial estimate - The EM procedure for dynamic state estimation with fixed shape (4.4.4) is used to provide an initial estimate  $\hat{\mathbf{x}}(t+1)$  for the iterations. It is seen that the ET-PMHT algorithm is more reactive to current measurements. A good shape estimate may degrade over time due to further

measurements. A gain reduction factor  $\alpha_{GR}$  is introduced, which reduces the cost of the initial estimate, and keeps the shape fixed.

2. Joint association and filtering - The EM procedure for general dynamic state estimation termed as ET-PMHT update (3.35) is used for jointly updating the shape and kinematics. Since the measurement function is unknown, the errors in the estimated measurement function propagate through the procedure. To reduce the effect of this issue, faces with an area below a threshold (small triangles) or faces with small angles (thin triangles) are not considered to be visible faces. Faces which are effectively associated to less than 3 measurements are also not considered to be visible faces to maintain the consistency of the linear system. The calculation of the updated covariance can be computationally expensive (size of innovation matrix). The measurement-to-face association with weights below a threshold can be dropped to speed up the covariance calculation.
3. Face management - The quality of each visible face  $m$  is calculated as  $Poisson(n^m, N_f^m)$  where  $N_f^m$  is calculated as (3.5) and  $n^m$  is the number of measurements effectively associated to the face.  $n^m$  is calculated as  $n^m = \sum_{r=1}^{n_t} w_t^{m,r}$ . If the quality of the face falls below a threshold  $\Gamma_{FM}$ , the face is deleted from the estimate.
4. Convex hull constraint - The convex hull operation is used on the estimate and the resulting convex hull is simplified.
5. Validation and face initialization - The measurements are validated with the visible parts of the resulting estimate. The measurements falling outside the

validation region of the visible faces are grouped by the edge of the visible faces they are closest to. For each of the groups, a new face is initialized using the edge and the farthest measurement to the edge in the group.

6. Convex hull and backwards propagation - The convex hull and simplification operation is performed since the estimate may not be convex after the addition of new faces. The resulting estimate is propagated backwards in time and is used as the initial estimate  $\hat{\mathbf{x}}(t)$  for the next iteration.

The update pseudocode is given in Algorithm 6. ET-PMHT-FS (4.4.4) is the update step with the shape being fixed. The variables with subscript FS are the output of the update when the shape is kept fixed.

---

**Algorithm 4** Update

---

```

procedure UPDATE( $Z(t + 1), \hat{\mathbf{x}}(t), P(t), \hat{X}_{DT}(t), \text{iter\_max}$ )
   $x^{FS}, P^{FS}, X_{DT}^{FS}, C^{FS} \leftarrow$  ET-PMHT-FS( $\hat{\mathbf{x}}(t), P(t), \hat{X}_{DT}(t), Z(t + 1)$ )
   $L_X, L_P, L_{DT}, L_C \leftarrow x^{FS}, P^{FS}, X_{DT}^{FS}, \alpha_{GR}, C^{FS}$     $x^0, P^0, X_{DT}^0 \leftarrow x^{FS}, P^{FS}, X_{DT}^{FS}$ 
  for  $l=1, l < \text{iter\_max}$  do
     $x^l, P^l, X_{DT}^l, C^l \leftarrow$  ET-PMHT-update( $x^{l-1}, P^{l-1}, X_{DT}^{l-1}, Z(t + 1)$ )
     $x^l, P^l, X_{DT}^l \leftarrow$  face_management( $x^l, P^l, X_{DT}^l$ )
     $x^l, P^l, X_{DT}^l \leftarrow$  simplified_convex_hull( $x^l, P^l, X_{DT}^l$ )
     $L_X, L_P, L_{DT}, L_C \leftarrow [L_X \ x^l], [L_P \ P^l], [L_{DT} \ X_{DT}^l], [L_C \ C^l]$ 
     $x^l, P^l, X_{DT}^l \leftarrow$  validate_and_initialize_faces( $x^l, P^l, X_{DT}^l, Z(t + 1)$ )
     $x^l, P^l, X_{DT}^l \leftarrow$  simplified_convex_hull( $x^l, P^l, X_{DT}^l$ )
     $x^l, P^l \leftarrow$  backpropagate( $x^l, P^l$ )
   $\hat{\mathbf{x}}(t + 1), P(t + 1), \hat{X}_{DT}(t + 1) \leftarrow$  min_cost_estimate( $L_X, L_P, L_{DT}, L_C$ )

```

---

## 3.6 Results

The performance of the algorithm is tested in various scenarios and the following metrics are used to evaluate the performance of the algorithm -

- RMSE of the center - The center of a convex shape is calculated as the mean of the vertices of the shape. Since both the target and the estimate are convex, their centers are calculated in the same manner as mentioned and denoted as  $c_T(t)$  and  $c_E(t)$ . The error in the center for each time frame  $t$ ,  $e_C(t)$  is calculated as

$$e_C(t) = \sqrt{|c_T(t) - c_E(t)|^2} \quad (3.43)$$

- RMSE of the velocity - The error in the velocity is calculated in a similar way as the center

$$e_V(t) = \sqrt{|v_T(t) - v_E(t)|^2} \quad (3.44)$$

- Intersection over Union IoU - This metric is a measure of the similarity of two shapes calculated for shapes  $A$  and  $B$  as follows

$$\text{IoU} = \frac{\text{Volume}(\text{Intersection of the shapes})}{\text{Volume}(\text{Union of the shapes})} \quad (3.45)$$

The intersection of two convex shapes is convex, however the union may not be convex. The IoU metric is thus approximated as follows

$$\text{IoU} = \frac{\text{Volume}(\text{Intersection of the shapes})}{\text{Volume}(\text{Convex Hull}(\text{Union of the shapes}))} \quad (3.46)$$

It must be noted that the approximation becomes better as the estimate is close to the target shape.

The whole target may not be observable over time and, hence the RMSE center and the IoU metrics need to be interpreted accordingly.

The performance of the proposed algorithm is compared with 3D Gaussian Process described in [8]. The algorithms are tested under the following scenarios

- Different levels of measurement noise when all parts of the target are visible
- Different levels of measurement noise when the target is only partially visible
- Different levels of number of measurements per unit area when the target is partially visible

### **3.6.1 3D Gaussian Process**

The extent of the target in this approach is described using a radial function which follows a Gaussian Process

$$f(\theta, \phi) \sim \mathcal{GP}(\mu(\theta, \phi), \kappa(\gamma, \gamma')) \quad (3.47)$$

where  $\gamma = (\theta, \phi)$ . The covariance function  $\kappa(\gamma, \gamma')$  describes the relation between different parts of the target given by

$$\kappa(\gamma, \gamma') = \sigma_f^2 e^{-\frac{d^2(\gamma, \gamma')}{2l^2}} + \sigma_r^2 \quad (3.48)$$

$$d(\gamma, \gamma') = \cos^{-1} \left( \cos(\phi) \cos(\phi') \cos(\theta) \cos(\theta') + \cos(\phi) \cos(\phi') \sin(\theta) \sin(\theta') + \sin(\phi) \sin(\phi') \right) \quad (3.49)$$

where  $l$  is the length scale,  $\sigma_f^2$  is the prior variance and  $d(\gamma, \gamma')$  is the distance function between two angle pairs. The process model is given by

$$\mathbf{x}(t+1) = F(t)\mathbf{x}(t) + \nu(t) \sim \mathcal{N}(0, Q(t)) \quad \mathbf{x}(t) = \begin{bmatrix} x_t(t) & x_r(t) & f(t) \end{bmatrix} \quad (3.50)$$

$x_t(t)$  denotes the translation component of the state with  $c(t)$  as the center,  $x_r(t)$  denotes the rotational component of the state and  $f(t)$  encodes the shape as a radial function. The measurement model is given by

$$z_{tl} = \tilde{h}(\mathbf{x}(t), z_{tl}) + e_{tl} \sim \mathcal{N}(0, R_{tl}) \quad (3.51)$$

$$\tilde{h}(\mathbf{x}(t), z_{tl}) = c(t) + p_{tl} H^f(\gamma_{tl}(c(t), q_t, z_{tl})) \quad (3.52)$$

$$H^f(\gamma_{tl}) = \kappa(\gamma_{tl}, \gamma_f) \kappa(\gamma_f, \gamma_f)^{-1} \quad (3.53)$$

$$R_{tl} = p_{tl} R_{tl}^f p_{tl}^T + R \quad (3.54)$$

Using initial estimate as  $\hat{\mathbf{x}}(0) = \mathcal{N}(\bar{\mathbf{x}}(0), P(0))$ , an EKF is used to update the state.

### 3.6.2 Scenario

The canonical scenario is such that the target is only partially visible to the sensor over time. The sensor is located at  $\begin{bmatrix} 0m & 0m & 0m \end{bmatrix}^T$  and the target is initially centered at  $\begin{bmatrix} 40m & 30m & 10m \end{bmatrix}^T$ . The target's initial velocity is  $\begin{bmatrix} -2ms^{-1} & -1ms^{-1} & -0.1ms^{-1} \end{bmatrix}^T$ . The target shape is shown in Figure [3.5] and the target trajectory is such that only the clear faces of the target are visible and the shaded faces are not visible throughout the entire run. 100 Monte Carlo runs are used to generate the results.

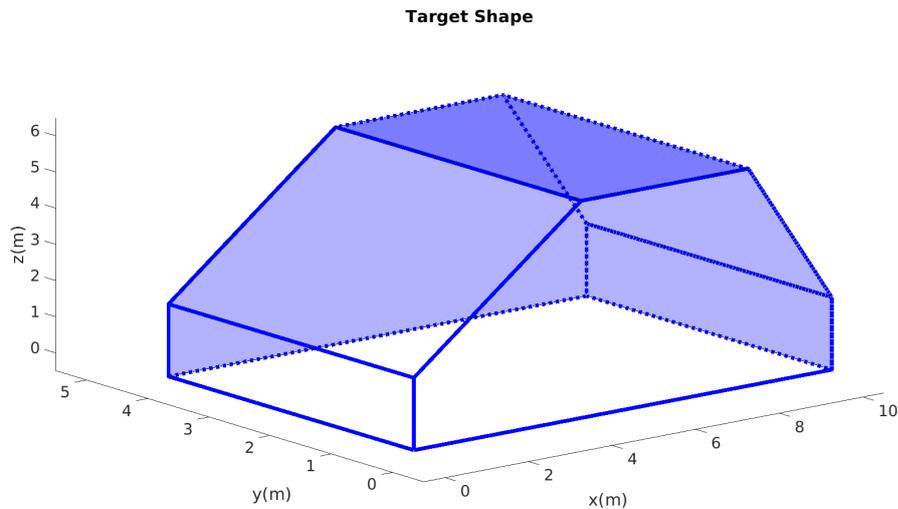


FIGURE 3.5: Target shape

The power spectral density of the process noise  $q$ , the measurement noise covariance and the average number of measurements per unit area are given by

$$q = 0.06m^2s^{-3} \quad R = \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{bmatrix} \quad \rho = \rho_N$$

The tracker parameters are

- Threshold for validation of measurements  $\Gamma_G = 0.99$
- Maximum number of iterations for EM - 10 for initialization and 5 for update.
- Threshold for exiting EM step -  $\tau_I = 0.1, \tau_U = 1$
- Threshold for simplifying the convex hull -  $\Gamma_S = 0.99$
- Threshold for face management -  $\Gamma_{FM} = 0.05$
- Gain reduction factor -  $\alpha_{GR} = 0.6$

Parameters for Gaussian Process - 642 radial basis points are chosen evenly spaced on a sphere. The parameters for the kernel are  $\sigma_f = 1m$ ,  $\sigma_r = 0.2m$  and  $l = \frac{\pi}{8}$ . The center is initialized with the mean of the initial measurements, velocity and the radii for the basis points are initialized to zeros.

### **3.6.3 Different levels of measurement noise when all parts of the target are visible**

The algorithms are tested for different levels of measurement noise while keeping the  $\rho_N = 0.5m^{-2}$  fixed. The standard deviation for the covariances used are  $\sigma_x, \sigma_y, \sigma_z = \{0.1m, 0.15m, 0.2m\}$ . It is assumed that all parts of the target are visible and generate measurements for this set of simulations. The Gaussian process does not account for self-occlusion and hence the scenario when the whole target is visible is within the assumptions of the GP algorithm.

Figures [3.6],[3.7],[3.8] show the results for the RMSE of the center and velocity and IoU. It can be seen that change in the measurement noise covariance has minimal effect on the RMSE velocity estimate. The RMSE center performance is better with lower measurement noise covariance. It must be noted that the center is not explicitly estimated as a part of the target state but is calculated from the vertices of the shape estimate. Hence, the variation in the RMSE of the center of the target needs to be interpreted accordingly. The IoU performance is clearly affected by the change in the measurement noise covariance and the IoU increases with the decrease in the measurement noise covariance. The proposed algorithm outperforms the Gaussian process in all three metrics for a convex polytope shape target even when all faces are visible since the ET-PMHT model is more appropriate for such targets.

The snapshots for a sample run is given with the ground truth, measurements, Gaussian Process estimate and the estimate of ET-PMHT for different times in

[3.9]. It is seen that the Gaussian Process estimate performs poorly compared to the proposed algorithm in terms of the shape. The faces of the target with a lesser area produce lesser number of measurements on an average. The poor performance is possibly due to the uneven distribution of measurements since the large faces are well estimated by the Gaussian Process.

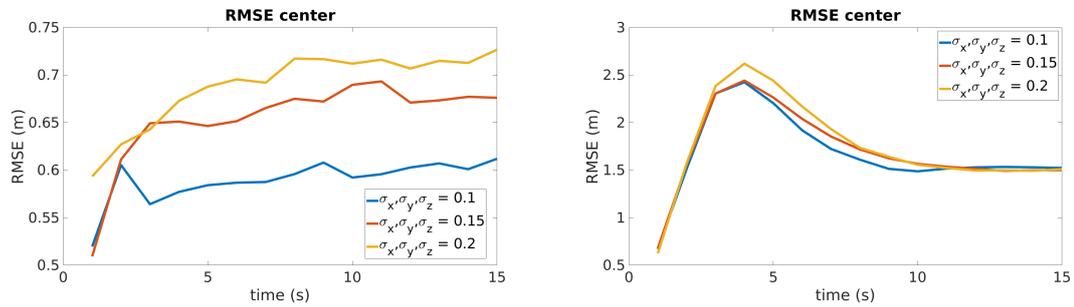


FIGURE 3.6: RMSE center for different levels of measurement noise when all faces are visible. Results for the proposed approach are on the left and the GP results are on the right

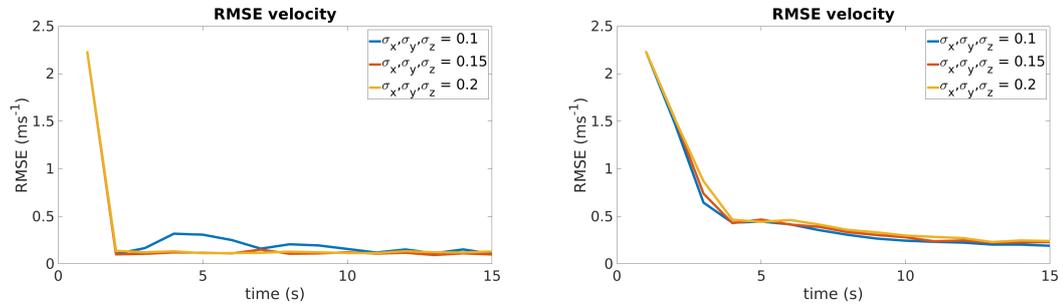


FIGURE 3.7: RMSE velocity for different levels of measurement noise when all faces are visible. Results for the proposed approach are on the left and the GP results are on the right

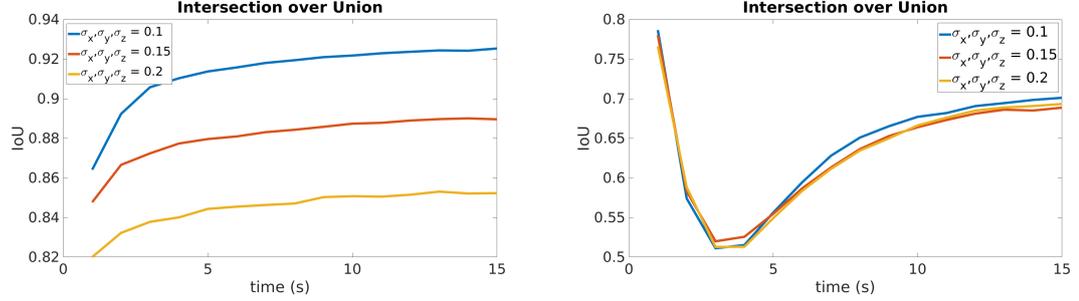


FIGURE 3.8: IoU results for different levels of measurement noise when all faces are visible. Results for the proposed approach are on the left and the GP results are on the right

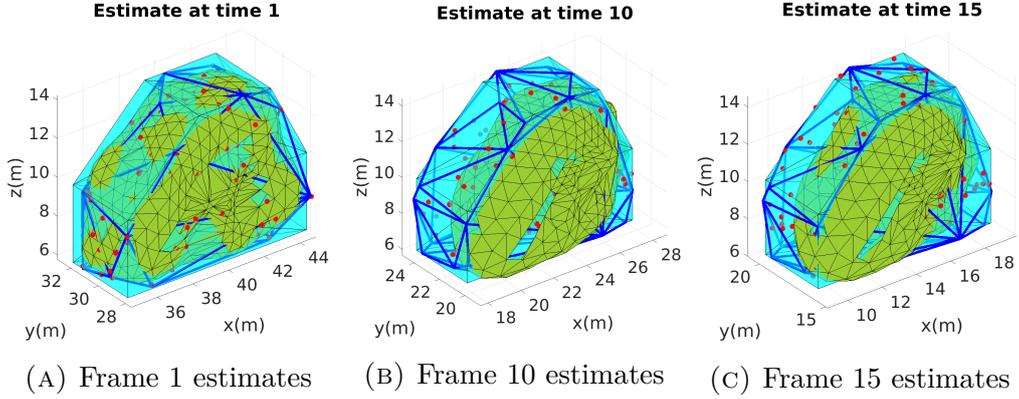


FIGURE 3.9: Estimates of the target shape at different times when all faces are visible for  $\sigma_x = \sigma_y = \sigma_z = 0.1m$ ,  $\rho_N = 0.5m^{-2}$ . The true target shape is given in cyan, the measurements at the current time are given in red. The estimate using the proposed algorithm is in blue and the GP estimate is in green.

### 3.6.4 Different levels of measurement noise with partial visibility

The algorithms are tested for different levels of measurement noise while keeping the  $\rho_N = 1m^{-2}$  fixed and the target is partially visible as described in 3.6.2. The standard deviation for the covariances used are  $\sigma_x = \sigma_y = \sigma_z$  which take the values

$\{0.1m, 0.15m, 0.2m\}$ .

Figures [3.10],[3.11],[3.12] show the results for the RMSE of the center and velocity and IoU. It must be noted that the RMSE center and IoU are calculated using only the parts of the target which were visible to the sensor. Snapshots of the ground truth, measurements, Gaussian Process estimate and ET-PMHT estimate for different times of a sample run are given in [3.13]. It is seen from the snapshots that the ET-PMHT is able to effectively estimate the shape even when the target is only partially visible, while the Gaussian Process estimate is not able to handle the issue of partial visibility. The GP estimate is close to the target for the faces which are visible but is not able to handle the faces which are not visible and hence, do not generate any measurements.

It can be seen from Figure 3.10 that the RMSE center performance for  $\sigma_x, \sigma_y, \sigma_z = 0.2m^{-1}$  is better than for  $\sigma_x, \sigma_y, \sigma_z = 0.1m^{-1}, 0.15m^{-1}$ . This anomaly occurs since the center is calculated from the shape estimate and not directly. A larger shape estimate may lead to a better center estimate than a more refined shape estimate. Clear effect of the change in measurement noise covariance can be seen from the IoU performance. As in the previous case of total visibility, the RMSE velocity is not affected to a significant extent by a change in the measurement noise covariance but the IoU shows increase in performance with the decrease in the measurement noise covariance. The IoU does not reach its maximum value for the ET-PMHT since its estimate can be biased for the faces which are not visible. A larger extent in the direction where the target is not visible has similar likelihood as the actual target shape and hence, the ET-PMHT estimate is usually larger in extent than the actual target especially for the parts which are not visible.

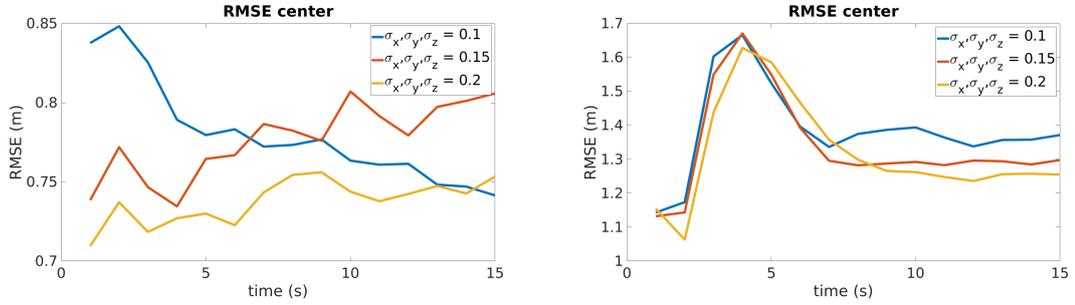


FIGURE 3.10: RMSE center for different levels of measurement noise with partial visibility. Results for the proposed approach are on the left and the GP results are on the right

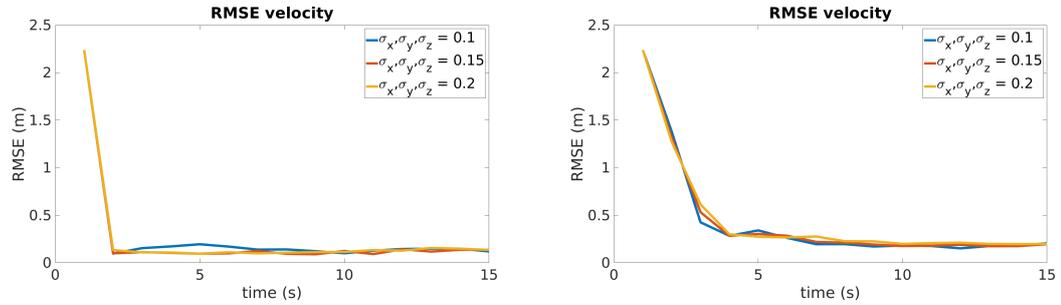


FIGURE 3.11: RMSE velocity for different levels of measurement noise with partial visibility. Results for the proposed approach are on the left and the GP results are on the right

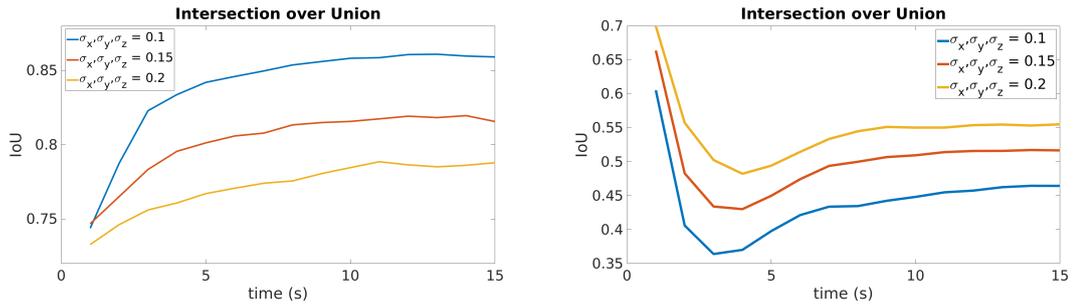


FIGURE 3.12: IoU results for different levels of measurement noise with partial visibility. Results for the proposed approach are on the left and the GP results are on the right

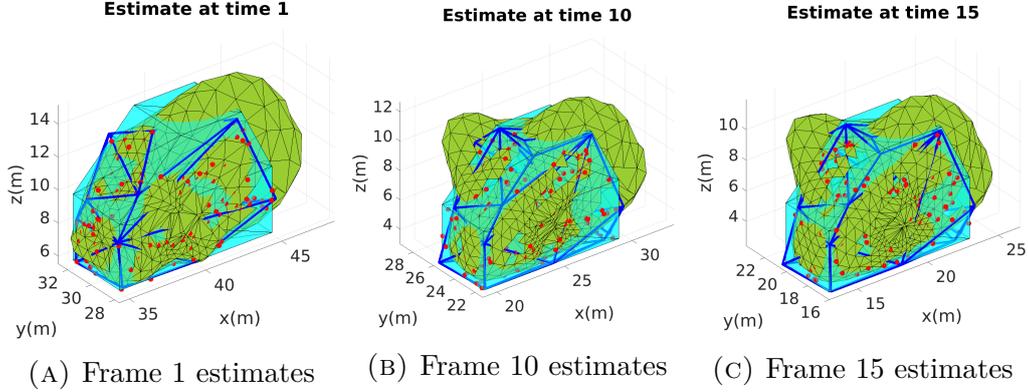


FIGURE 3.13: Estimates of the target shape at different times  $\sigma_x = \sigma_y = \sigma_z = 0.1m$ ,  $\rho_N = 0.1m^{-2}$ . The true target shape is given in cyan, the measurements at the current time are given in red. The estimate using proposed algorithm is in blue and the GP estimate is in green

### 3.6.5 Different levels of average number of measurements per unit area with partial visibility

The average number of measurements per unit area is varied as  $\rho_N = \{0.5m^{-2}, 1m^{-2}, 1.5m^{-2}\}$  while keeping the measurement noise covariance fixed as  $\sigma_w = 0.1m$  for  $w = x, y, z$ .

Figures [3.14],[3.15],[3.16] show the RMSE of the center, velocity and IoU for different levels of measurements when the target is only partially visible. Similar to the previous scenarios, the effect of the different levels of measurements is not significant for RMSE velocity. The RMSE center performance is better with increase in the number of measurements generated per unit area. The IoU performance is significantly better when  $\rho_N$  changes from a low value of  $1m^{-2}$  to  $1.5m^{-2}$  but the difference between  $\rho_N = 1.5m^{-2}$  and  $\rho_N = 2m^{-2}$  is not significant. Hence,

increasing the density of the measurement above certain level does not affect the IoU significantly.

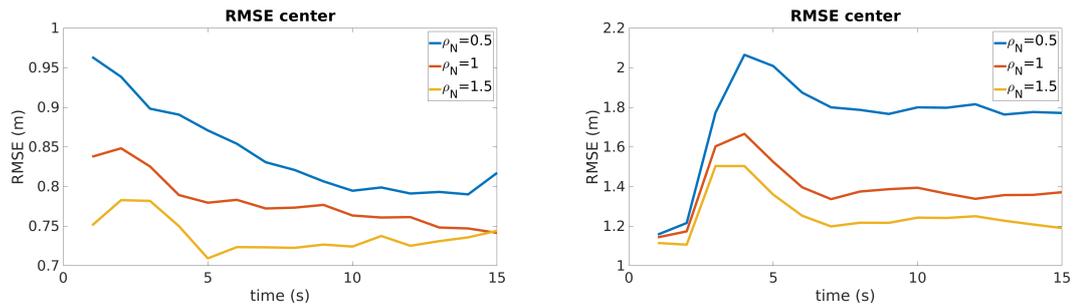


FIGURE 3.14: RMSE center for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right

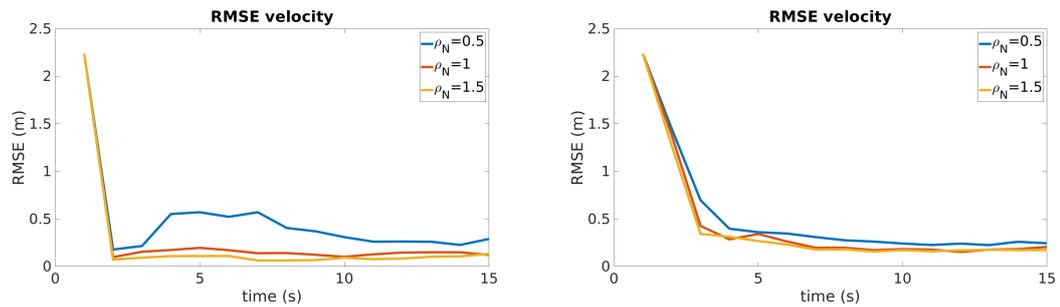


FIGURE 3.15: RMSE velocity for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right

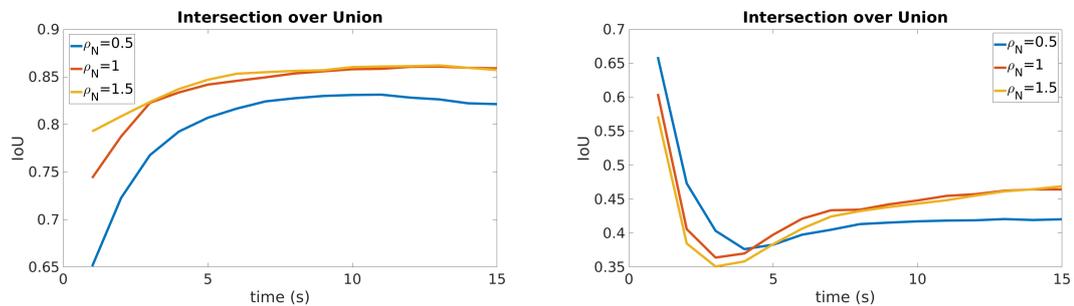


FIGURE 3.16: IoU results for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right

### **3.7 Conclusions and Future work**

The problem of tracking a 3D extended target with convex polytope shape even when the target is partially visible was discussed. A convex polytope model based on the vertices and Delaunay triangulation was used and a linear measurement model was developed for the shape. Different parts of the extended targets were treated as separate targets constrained by the rigid body dynamics of the target and an algorithm was proposed under the MTT framework [16]. In particular, the point target PMHT was modified and extended target PMHT (ET-PMHT) equations (3.20),(3.35),(4.4.4) were derived. An algorithm for initialization and update was developed using ET-PMHT joint association and filtering. The proposed algorithm was compared with 3D Gaussian Process [8] for RMSE center, velocity and IoU performance metrics.

Some of the issues to be addressed in future works are discussed. One of the limitations of the current algorithm considered is the restriction to convex shapes. While the ET-PMHT filtering supports non-convex shapes as well, developing a complete algorithm for tracking 3D non-convex shapes can be included in future work. The current algorithm becomes biased when the parts of the target which were previously not visible become visible over time. One approach to handling this issue is to label each face of the estimate and use more advanced track management to obtain better performance. An extension of the current work to address the problem of clutter and multiple extended targets can also be considered for future work.

### 3.8 Appendix - Derivatives of the auxiliary function

The derivatives of the auxiliary function  $\mathcal{L}(\mathbb{X}|\hat{\mathbb{X}})$  (3.31) with respect to the state  $x_t$  are given below -

$$\begin{aligned}
 \nabla_{\mathbf{x}(t)|t \neq 0, T} \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}}) &= 0 \\
 \implies \left[ - \sum_{m=1}^M (Q_{t-1}^m)^{-1} F \right] \mathbf{x}(t-1) \\
 + \left[ \sum_{m \in M_t^{vis}} (H_t^m F^{mx})^T (R_t^m)^{-1} (H_t^m F^{mx}) + \sum_{m=1}^M (Q_{t-1}^m)^{-1} + F^T (Q_t^m)^{-1} F \right] \mathbf{x}(t) \\
 + \left[ - \sum_{m=1}^M F^T (Q_t^m)^{-1} \right] \mathbf{x}(t+1) &= \left[ \sum_{m \in M_t^{vis}} (H_t^m F^{mx})^T (R_t^m)^{-1} \right] Z_t \quad (3.55)
 \end{aligned}$$

$$\begin{aligned}
 \nabla_{\mathbf{x}(0)} \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}}) &= 0 \\
 \implies \left[ \sum_{m=1}^M (P_0^m)^{-1} + F^T (Q_0^m)^{-1} F \right] \mathbf{x}(0) + \left[ - F^T (Q_0^m)^{-1} \right] \mathbf{x}(1) \\
 = \left[ \sum_{m=1}^M (P_0^m)^{-1} \right] \bar{\mathbf{x}}(0) \quad (3.56)
 \end{aligned}$$

$$\begin{aligned}
 \nabla_{\mathbf{x}(T)} \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}}) &= 0 \\
 \implies \left[ \sum_{m=1}^M -(Q_{T-1}^m)^{-1} F \right] \mathbf{x}(T-1) \\
 + \left[ \sum_{m \in M_T^{vis}} (H_T^m F^{mx})^T (R_T^m)^{-1} (H_T^m F^{mx}) + \sum_{m=1}^M (Q_{T-1}^m)^{-1} \right] \mathbf{x}(T) \\
 = \left[ \sum_{m \in M_T^{vis}} (H_T^m F^{mx})^T (R_T^m)^{-1} \right] Z_T \quad (3.57)
 \end{aligned}$$

## Bibliography

- [1] *Convex Hulls*, pages 243–258. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-77974-2. doi: 10.1007/978-3-540-77974-2\_11. URL [https://doi.org/10.1007/978-3-540-77974-2\\_11](https://doi.org/10.1007/978-3-540-77974-2_11).
- [2] Marcus Baum and Uwe Hanebeck. Tracking an extended object modeled as an axis-aligned rectangle. pages 2422–2427, 01 2009.
- [3] Shishan Yang and Marcus Baum. Tracking the orientation and axes lengths of an elliptical extended object. *IEEE Transactions on Signal Processing*, 67(18):4720–4729, 2019. doi: 10.1109/TSP.2019.2929462.
- [4] Johann Wolfgang Koch. Bayesian approach to extended object and cluster tracking using random matrices. *IEEE Transactions on Aerospace and Electronic Systems*, 44(3):1042–1059, 2008. doi: 10.1109/TAES.2008.4655362.
- [5] Jian Lan and X. Rong Li. Tracking of maneuvering non-ellipsoidal extended object or target group using random matrix. *IEEE Transactions on Signal Processing*, 62(9):2450–2463, 2014. doi: 10.1109/TSP.2014.2309561.
- [6] Ben Liu, Ratnasingham Tharmarasa, Rahim Jassemi, Daly Brown, and Thia Kirubarajan. Extended target tracking with multipath detections, terrain-constrained motion model and clutter. *IEEE Transactions on Intelligent Transportation Systems*, 22(11):7056–7072, 2021. doi: 10.1109/TITS.2020.3001174.

## BIBLIOGRAPHY

---

- [7] Niklas Wahlstrom and Emre Ozkan. Extended target tracking using gaussian processes. *IEEE Transactions on Signal Processing*, 63(16):4165–4178, 2015. doi: 10.1109/TSP.2015.2424194.
- [8] Murat Kumru and Emre Ozkan. Three-dimensional extended object tracking and shape learning using gaussian processes. *IEEE Transactions on Aerospace and Electronic Systems*, 57(5):2795–2814, oct 2021. doi: 10.1109/taes.2021.3067668. URL <https://doi.org/10.1109/2Ftaes.2021.3067668>.
- [9] Xu Tang, Mingyan Li, Ratnasingham Tharmarasa, and Thiagalingam Kirubarajan. Seamless tracking of apparent point and extended targets using gaussian process pmht. *IEEE Transactions on Signal Processing*, 67(18):4825–4838, 2019. doi: 10.1109/TSP.2019.2932873.
- [10] Yunfei Guo, Yong Li, Ratnasingham Tharmarasa, Thiagalingam Kirubarajan, Murat Efe, and Bahadir Sarikaya. Gp - pda filter for extended target tracking with measurement origin uncertainty. *IEEE Transactions on Aerospace and Electronic Systems*, 55(4):1725–1742, 2019. doi: 10.1109/TAES.2018.2875555.
- [11] Marcus Baum and Uwe D. Hanebeck. Shape tracking of extended objects and group targets with star-convex rhms. In *14th International Conference on Information Fusion*, pages 1–8, 2011.
- [12] Antonio Zea, Florian Faion, Marcus Baum, and Uwe D. Hanebeck. Level-set random hypersurface models for tracking nonconvex extended objects. *IEEE Transactions on Aerospace and Electronic Systems*, 52(6):2990–3007, 2016. doi: 10.1109/TAES.2016.130704.

## BIBLIOGRAPHY

---

- [13] Abdullahi Daniyan, Sangarapillai Lambotharan, Anastasios Deligiannis, Yu Gong, and Wen-Hua Chen. Bayesian multiple extended target tracking using labeled random finite sets and splines. *IEEE Transactions on Signal Processing*, 66(22):6076–6091, 2018.
- [14] Hauke Kaulbersch, Jens Honer, and Marcus Baum. A cartesian b-spline vehicle model for extended object tracking. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 1–5, 2018. doi: 10.23919/ICIF.2018.8455717.
- [15] Karl Granstrom, Marcus Baum, and Stephan Reuter. Extended object tracking: Introduction, overview and applications, 2017.
- [16] Prabhanjan Mannari, Ratnasingham Tharmarasa, and Thiagalingam Kirubarajan. Extended target tracking under multitarget tracking framework for convex polytope shapes. *preprint on SSRN*, 2022. doi: <http://dx.doi.org/10.2139/ssrn.4290822>.
- [17] S. S. Rao. *The finite element method in engineering*. Elsevier/Butterworth Heinemann, Amsterdam ;, 4th ed. edition, 2005. ISBN 1-280-96441-3.
- [18] Han X. Gaetjens Samuel J. Davey. *Track-Before-Detect Using Expectation Maximisation*. Springer Singapore. ISBN 978-981-13-3971-4. URL <https://link.springer.com/book/10.1007/978-981-10-7593-3>.

## Chapter 4

# ET-PMHT for Tracking 3D

# Extended Targets with

# Maneuvers and Clutter

The content of this chapter is ready to be submitted to a Journal as of December 8, 2023.

---

**Mannari, Prabhanjan** and Tharmarasa, Ratnasingham and Kirubarajan, Thiagalingam, "ET-PMHT for Tracking 3D Extended Targets with Maneuvers and Clutter"

---

# **ET-PMHT for Tracking 3D Extended Targets with Maneuvers and Clutter**

## **Abstract**

This article addresses the problem of extended target tracking of a 3D convex polytope shape with maneuvers and in the presence of clutter. Extended targets (as opposed to point targets) are those that occupy multiple sensor resolution cells. Hence, it is possible to obtain multiple measurements for the same frame from a single extended target and the shape as well as the kinematics of the target need to be estimated using measurements over time frames. In the current work, different parts of the extended target are treated as distinct targets constrained by the target rigid body shape, and an algorithm is developed under the extended target tracking framework proposed in our previous works. ET-PMHT equations for a maneuvering target are derived by linearizing the transition function to handle the nonlinearity of the maneuvers. The visibility problem of self-occlusion is handled by associating the current measurements to the visible faces of the target. Clutter is incorporated into the ET-PMHT and face management is used to delete erroneous faces. The algorithm is able to handle clutter that occurs from the interior of the target as well. The algorithm supports adding new faces to the estimate as faces of the target that were previously not visible become visible over time. The performance of the algorithm is compared with the 3D Gaussian Process for various scenarios of different measurement noise covariance, different levels of clutter density and different levels of density of measurements from the target. The RMSE

of the center, RMSE of the velocity and Intersection over Union (IoU) metrics are used to quantify the performance.

**Keywords:** *Extended Target, Self-Occlusion, Expectation Maximization, Probabilistic Multiple Hypothesis Tracker (PMHT), Hausdorff distance*

## 4.1 Introduction

Traditional target tracking algorithms assume ‘point’ targets. Point targets are those which have size/extent such that the entire target occupies at most a single resolution cell of the sensor. Hence, these targets produce at most one measurement per frame. Recent advancements have led to wide availability of high resolution sensors, such as automotive radar and LiDAR (Light Detection and Ranging), for target tracking applications. Since the sensors have better resolution, the targets may occupy multiple sensor resolution cells and hence can generate multiple measurements in a single frame. Such targets are termed as extended targets, and have finite extent/size compared to the sensor resolution cell. The sensor resolution size can vary with the distance from the sensor and hence the same target can be a point target and an extended target depending on the relative position of the target with respect to the sensor. The measurements from the target contain information about the target shape, and hence it is necessary to estimate the shape as well as the kinematics of the target.

The primary challenge in tracking extended targets is to estimate their shape in the presence of joint uncertainty in the shape and kinematics, i.e., targets with

distinct shapes and kinematics can generate the same set of measurements over time. Yet another challenge in extended target tracking is the problem of self-occlusion. The measurements are generated only from the visible parts of the target. Generally, for an extended target, different parts of the target are visible at different times, and the entirety of the target is not visible at the same time. It may be possible that some parts of the target are not visible over the whole time duration. The objective is to estimate the entire shape or the shape of the observable parts of the target using such measurements. In a high resolution scenario, the measurements for the target are usually obtained by clustering to eliminate detections from the background such as ground clutter. However, some of the clutter points may get included in the target cluster as well. The presence of clutter can degrade the shape estimation significantly if not accounted for. The measurements may occur from the interior of the target as well and need to be handled appropriately.

Several approaches to extended target tracking have been proposed in the literature. In [1],[2],[3], simple shapes such as rectangles and ellipses are used to describe the target shape. The parameters of the shape such as length/breadth and major/minor axes are to be estimated. The Random Matrix approach used in [1] uses a symmetric positive definite (PSD) matrix to describe the extent of an extended target or a group of targets moving in a coordinated manner. More specifically, the matrix determines the covariance of the measurements obtained from the target around its center. An analytical Bayesian solution is derived by assuming that the extent matrix is Inverse Wishart distributed. The Random Matrix shape description is appropriate for ellipsoidal shapes but fails to capture

more complicated features of the target shape. In [4], the Random Matrix model is extended to describe non-ellipsoidal shapes using multiple ellipsoids. Multipath detections and clutter issues are considered in a terrain constrained environment using the Random Matrix model in [5].

The Gaussian Process (GP) model is introduced for extended target tracking in [6] and extended to 3D in [7]. The shape is described using a radial basis function that follows a Gaussian Process. This star-convex model can handle a variety of complicated shapes and their features. An appropriate kernel function is used for the covariance between the basis points, with the squared-exponential function being used generally. An Extended Kalman Filter (EKF) is used to handle the nonlinearity and an analytical expression is derived for the Jacobian required for filtering. Further details of the Gaussian Process are presented in Section 4.6.1. The article [8] addresses the issue when the same target changes from an extended target to a point target (or vice-versa) due to the relative change in the sensor-target geometry, using a Poisson rate for the number of measurements from the target. Clutter is included in the GP model using Probabilistic Data Association (PDA) in article [9].

The Random Hypersurface Model described in [10] can be used for a variety of simple and complicated shapes. The model supports simple shapes such as ellipses as well as more complicated shapes using a polygonal shape descriptor. A pseudo-measurement equation is formed, which describes the distances from the shape to the measurements, with uncertainty in the measurement source location. The spline model described in [11] is yet another model to describe the target shape. A detailed overview of extended target tracking techniques is presented in [12].

The current work extends our previous works [13],[14] to track 3D convex polytope shaped targets with maneuvers and in the presence of clutter. The framework described in [13] is used and different faces of the target are treated as distinct targets constrained by the rigid body target shape. This allows the determination of the faces that are visible and association of measurements only to those faces to handle the problem of self-occlusion, which is not sufficiently addressed in the literature. The center of the target is generally a part of the state to be estimated in existing works. However, the true center of the target may not be observable even with measurements over time, particularly when the target is only partially visible. In [13], these problems have been addressed for a 2D convex polytope shape and Nearly Constant Velocity (NCV) model. The extended target PMHT (ET-PMHT) proposed in [14] is able to track a 3D convex polytope shaped target with the NCV model even when the target is only partially visible. However, these works do not consider target maneuvers, which is common in reality. Clutter and measurements from the interior of the target are not supported as well. The current work proposes to relax these assumptions and develop an algorithm for a more realistic scenario.

In the current work, the center of the target is not a part of the state, and the shape is described by the boundary alone. The proposed algorithm is able to handle clutter and measurements occurring from the interior of the target. The number of faces used to describe the shape can be altered using face management to delete erroneous faces and face initialization to add new faces, as opposed to existing approaches where a fixed number of parameters are used to describe the shape. This is particularly useful for adding new faces when new parts of the

target become visible over time or to delete erroneous faces created due to clutter.

In particular, the target shape is described by a convex hull and its Delaunay triangulation. Describing arbitrary shapes using triangular meshes is a technique widely used in engineering. Different faces of the target are then treated as distinct targets constrained by the triangulation, and an ET-PMHT is derived for joint association and filtering. The nonlinearity of the maneuvering is handled using a Coordinated Turn (CT) model for the target dynamics and linearization for filtering. The clutter is incorporated into the ET-PMHT by reweighting the measurement to face association weights with clutter probability. Face Management is used to delete faces with low quality, i.e. the actual number of measurements associated to the face is low compared to the expected number. Face-to-face association is developed using the Hausdorff distance to associate faces of the estimate across time frames. Face Initialization is used to add new faces using measurements outside the validation region of the shape.

The main contributions of this paper are

- A 3D convex polytope model with rotational dynamics to handle maneuvers was developed.
- ET-PMHT equations were derived using linearization to handle the nonlinearity of the dynamics model.
- The visibility problem of self-occlusion is addressed.
- The algorithm is able to handle clutter and measurements occurring from the interior of the target.

- Face-to-face association is proposed using Hausdorff distance as a suitable distance measure between the face triangulations to track faces across frames.
- Simulations of various scenarios were performed to evaluate the performance of the proposed algorithm and compare the performance with the Gaussian Process for different metrics.

The tracking problem is described in Section 4.2 and some of the preliminary aspects of the solution are given in Section 4.3. The equations for the ET-PMHT are derived in Section 4.4 with the algorithm developed in Section 4.5. The results are given in Section 4.6 and the conclusions are presented in Section 4.7.

## **4.2 Problem Description**

The article discusses the problem of tracking a single extended target with known dynamics in 3D. The target is a rigid body with convex polytope shape. The target can exhibit maneuvers by rotation about the  $z$  axis, i.e. yaw rotation. The shape as well as the kinematics of the target are to be estimated using the measurements from the target over time. Measurements occur only from the visible parts of the target and not the entire target due to self-occlusion. The effect of clutter is included in the scenario, including those from the interior of the target.

### 4.2.1 Target Model

The target model at time frame  $t$  consists of state  $\mathbf{x}(t)$ , which includes the shape descriptor  $X_S(t)$  and the kinematics  $v(t)$  and  $\Omega(t)$ , where  $v(t)$  is the translational velocity and  $\Omega(t)$  is the turn rate. It is assumed that the target only undergoes rotation about the  $z$ -axis (yaw rotation). The model supports the case where the target is maneuvering  $\Omega(t) \neq 0$  as well as NCV model as  $\Omega(t) \rightarrow 0$ . The target shape is represented by the shape descriptor  $X_S(t)$  and a Delaunay triangulation matrix  $X_{DT}$ , which divides the surface of the target into non-overlapping triangles. The shape descriptor  $X_S(t)$  consists of  $N_X$  vertices and the edges between the vertices are described in the triangulation. Each row of the triangulation consists of the indices of the vertices that make up the triangular face. The target is a rigid body, and hence the triangulation does not change with the target motion. As such, it is not explicitly included in the target state. The target state can be explicitly written as

$$\begin{aligned} \mathbf{x} &= \left[ X_S^T \quad v^T(t) \quad \Omega(t) \right]^T & (4.1) \\ X_S(t) &= \left[ p_1^T(t) \quad p_2^T(t) \quad \dots \quad p_{N_X}^T(t) \right]^T \text{ with} \\ p_i^T(t) &= \left[ p_{ix}(t) \quad p_{iy}(t) \quad p_{iz}(t) \right] \quad v^T(t) = \left[ v_x(t) \quad v_y(t) \quad v_z(t) \right] \end{aligned}$$

Let  $T$  be the time between time frames  $t$  and  $t + 1$ . The target dynamics are given by -

$$\mathbf{x}(t + 1) = c^{Nx}(t + 1) + F_R^{Nx}(t) \left[ F_T^{Nx}(t) \mathbf{x}(t) - c^{Nx}(t + 1) \right] + \Gamma_4^{Nx} \nu(t) \quad (4.2)$$

$$= f_t(\mathbf{x}(t)) + \Gamma_4^{Nx} \nu(t) \quad (4.3)$$

where  $\nu(t)$  is the process noise with process noise covariance  $Q$ . The superscripts on the variables  $c, F_R, F_T, \Gamma_4$  are used to indicate the number of vertices in the vector the variables are operated with. This allows for overloading of the same notation when used with vectors of different sizes. The target dynamic matrices are given by -

$$F_T^{Nx} = \begin{bmatrix} I_3 & 0 & 0 & \dots & F_{pv} & 0 \\ 0 & I_3 & 0 & \dots & F_{pv} & 0 \\ 0 & 0 & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_3 & F_{pv} & 0 \\ 0 & 0 & \dots & 0 & F_{vv} & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix} \quad F_R^{Nx} = \begin{bmatrix} F_r & 0 & 0 & \dots & 0 \\ 0 & F_r & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & F_r & 0 \\ 0 & 0 & \dots & 0 & I_4 \end{bmatrix}$$

$$F_r = \begin{bmatrix} \cos(\Omega(t)T) & -\sin(\Omega(t)T) & 0 \\ \sin(\Omega(t)T) & \cos(\Omega(t)T) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$F_{pv} = \begin{bmatrix} \frac{\sin(\Omega(t)T)}{\Omega(t)} & -\frac{(1-\cos(\Omega(t)T))}{\Omega(t)} & 0 \\ \frac{(1-\cos(\Omega(t)T))}{\Omega(t)} & \frac{\sin(\Omega(t)T)}{\Omega(t)} & 0 \\ 0 & 0 & T \end{bmatrix} \quad F_{vv} = \begin{bmatrix} \cos(\Omega(t)T) & -\sin(\Omega(t)T) & 0 \\ \sin(\Omega(t)T) & \cos(\Omega(t)T) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.4)$$

The matrix  $c^{N_x}(t + 1)$  is used to rotate the target about its translated center  $c(t + 1)$ , given by

$$c^{N_x}(t + 1) = \left[ c(t + 1)^T \quad c(t + 1)^T \quad \dots \quad c(t + 1)^T \quad 0_{1 \times 4} \right]^T \quad (4.5)$$

The translated center  $c(t + 1)$  is calculated as the mean of the vertices of the translated shape.

$$c(t + 1) = \bar{M}^{N_x} F_T^{N_x}(t) \mathbf{x}(t), \quad \text{where } \bar{M}^{N_x} = \frac{1}{N_x} \begin{bmatrix} I_3 & I_3 & \dots & I_3 & 0_{3 \times 4} \end{bmatrix} \quad (4.6)$$

The matrix  $\Gamma_4^{N_x}$  is used to add the same process noise to each vertex since the target is a rigid body and all the vertices move together. The process noise covariance  $Q$  and  $\Gamma_4^{N_x}$  are described below.

$$Q = \begin{bmatrix} q_x & 0 & 0 & 0 \\ 0 & q_y & 0 & 0 \\ 0 & 0 & q_z & 0 \\ 0 & 0 & 0 & q_\Omega \end{bmatrix} \quad \Gamma_4^{N_x} = \begin{bmatrix} \frac{1}{2}T^2 & 0 & 0 & 0 \\ 0 & \frac{1}{2}T^2 & 0 & 0 \\ 0 & 0 & \frac{1}{2}T^2 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ T & 0 & 0 & 0 \\ 0 & T & 0 & 0 \\ 0 & 0 & T & 0 \\ 0 & 0 & 0 & T \end{bmatrix} \quad (4.7)$$

## 4.2.2 Measurement Model

It is assumed that the measurements are generated from the surface of the target corrupted by zero mean Gaussian noise with known covariance. The measurements occur only from the visible parts of the target due to self-occlusion. For a convex polytope shape, a face of the target is visible to the sensor if the centroid of the shape and the location of the sensor are on the opposite sides of the face. The number of measurements from a face is Poisson distributed with average number of measurements  $n_{avg}$  proportional to the area of the face  $A$ .

$$n_{avg} = \rho \times A \tag{4.8}$$

The effect of the sensor target geometry can be incorporated into the average number of measurements by using an effective area  $A_e = A \times \sin(\frac{\alpha}{2})$ , where  $\alpha$  is the solid angle subtended by the face at the sensor. This ensures that a face far away from the sensor or an inclined face produces lesser number of measurements compared to the same face being closer to the sensor or facing the sensor directly. The effect of external conditions such as time of the day and weather can be included in the density of measurements by having a space and time dependent  $\rho(x, y, z, t)$ . This reflects a more realistic scenario where the density can be lowered in bad visibility conditions such as snow and fog. In the current work, as a first step towards the solution, Equation 4.8 is used directly without incorporating more realistic conditions.

Consider a target  $\mathbf{x}$  with  $M$  number of faces (i.e. number of rows in  $X_{DT}$

matrix), with each face represented by vertices  $(p_{i1}, p_{i2}, p_{i3})$  for  $i = 1$  to  $M$ . Let  $N_{z_i}$  measurements occur from the  $i^{\text{th}}$  face with total number of measurements being  $N_z$ . Each measurement  $z_{ij}$  from face  $i$  can be written as -

$$z_{ij} = \lambda_{ij1}p_{i1} + \lambda_{ij2}p_{i2} + \lambda_{ij3}p_{i3} + w \quad (4.9)$$

$$\text{where } w \sim \mathcal{N}(0, R) \quad \text{and } \lambda_{ijk} \geq 0 \quad \sum_{k=1}^3 \lambda_{ijk} = 1$$

$R$  denotes the measurement noise covariance.

Each measurement  $z_{ij}$  is a convex combination of the vertices of the corresponding face corrupted by a noise term. The association of the measurements to the target faces and the source locations of the measurements described by  $\lambda$ 's are unknown. The total measurement set can be written in terms of the state and a measurement matrix as shown below in Equation (4.10). The rows in the measurement matrix for each measurement consist of the coefficients for the vertices corresponding to the measurement and zero for other parameters. Measurements  $z_{11}, z_{12}$  occur from face 1 with vertices  $(p_1, p_2, p_3)$ , zero measurements occur from face 2 with vertices  $(p_2, p_3, p_4)$ , measurement  $z_{31}$  occurs from face 3 with vertices  $(p_1, p_4, p_6)$  and measurements  $z_{41}, z_{42}$  occur from face 4 with vertices  $(p_4, p_5, p_6)$ .

$$\begin{bmatrix} z_{11} \\ z_{12} \\ z_{31} \\ z_{41} \\ z_{42} \end{bmatrix} = \begin{bmatrix} \lambda_{111} & \lambda_{112} & \lambda_{113} & 0 & 0 & 0 \\ \lambda_{121} & \lambda_{122} & \lambda_{123} & 0 & 0 & 0 \\ \lambda_{311} & 0 & 0 & \lambda_{312} & 0 & \lambda_{313} \\ 0 & 0 & 0 & \lambda_{411} & \lambda_{412} & \lambda_{413} \\ 0 & 0 & 0 & \lambda_{421} & \lambda_{422} & \lambda_{423} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{bmatrix} + \begin{bmatrix} w_{11} \\ w_{12} \\ w_{31} \\ w_{41} \\ w_{42} \end{bmatrix} \quad (4.10)$$

The clutter is generated in a region around the target described by the cuboid  $C_C$  with volume  $V_C$ . The measurements occurring from the interior of the target do not provide shape information about the target, and hence, they are modeled in the current work using clutter occurring from inside the target. The clutter is generated with a clutter density  $\rho_C$  and the number of clutter points is Poisson distributed with average number of clutter points being  $\bar{N}_C$ .  $\bar{N}_C$  is calculated as -

$$\bar{N}_C = \rho_C V_C \quad (4.11)$$

The cuboid  $C_C$  is divided into cells with ranges  $C_x, C_y, C_z$  in the  $x, y, z$  directions, respectively. The clutter points are uniformly distributed across the cells of  $C_C$ . Hence, the probability of a measurement being part of clutter  $P_C$  is given by

$$P_C = \frac{\bar{N}_C}{C_x C_y C_z} \quad (4.12)$$

An example of a target with measurements and clutter is shown in Figure 4.1.

## **4.3 Solution**

### **4.3.1 Extended Target Tracking Framework**

In the previous works, [13] and [14], the Point Multitarget Tracking Framework was used to develop a single Extended Target Tracking Framework by assuming

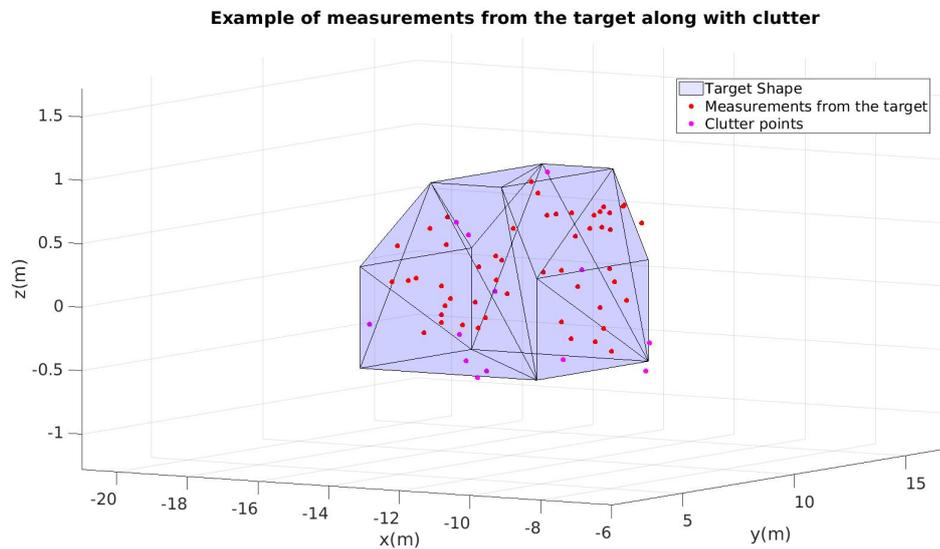


FIGURE 4.1: Example of measurements from the target along with clutter.

different parts of the extended target as multiple distinct targets. This framework is used in the current work to modify the existing point target Probabilistic Multiple Hypotheses Tracker (PMHT) to handle Extended Target Tracking with rotation and clutter. A brief description of the Point Multitarget Tracking Framework and its extension to the Extended Target Tracking Framework is discussed below. The Point Multitarget Tracking Framework consists of the following steps:

- Gating - The measurements at the current time step are validated using the existing tracks. Measurements that do not fall inside the validation region of any existing track are grouped separately.
- Association and Filtering - Measurements occurring in the validation region of a certain track are associated to that track. In case a measurement lies in the validation region of multiple tracks, multitarget association [15],[16]

needs to be used. Tracks are filtered using the measurements they are associated to. Multiple tracks are filtered jointly in case of common measurement association.

- Track Initialization - New tracks are initialized using the measurements that do not fall in the validation region of any existing track.
- Track Management - Existing tracks can be deleted and tentative tracks can be confirmed by using either logic or quality based track management [16]. Both metrics depend on the actual number of measurements received from the track versus the expected number of measurements from the track.

In the Extended Target Tracking Framework, different faces of the Extended Target are treated as different targets constrained by the rigid body motion of the target, i.e. the faces may have some points in common and all faces have the same kinematics. An Extended Target - Probabilistic Multiple Hypotheses Tracker (ET-PMHT) joint association and filtering is derived in Sections 4.4.1 and 4.4.3 under these assumptions to handle Extended Target tracking with rotation and in the presence of clutter. The measurements are validated using existing faces of the target and new faces are added for measurements outside the validation region of any existing face. This is particularly useful to handle occlusion as new faces of the target that were previously not visible become visible over time. Face Management is used to delete erroneous faces estimated, especially in the presence of clutter.

### 4.3.2 Determination of the measurement function

Since the source locations of the measurements are unknown, it becomes necessary to associate the measurements to a face and estimate the source location as in Equation (4.9). The source location is estimated independently for each face-measurement pair. The source location for measurement  $z^r$  from face  $f^m$  is the closest point in  $f^m$  to the measurement  $z^r$ . Since the face  $f^m$  is a triangle, the closest point calculation can be done using geometric methods. The source location  $\hat{z}^r$  can be written in terms of the face  $f^m$  as

$$\hat{z}^r = h^{m,r} f^m \quad (4.13)$$

It must be noted that even though  $h^{m,r} f^m$  yields the source location  $\hat{z}^r$ , it is the best possible estimate (closest in terms of distance) of the measurement function, and will be used as the estimate of the measurement function itself in later sections. The total set of measurements can be written in vector form as

$$\begin{bmatrix} z^1 \\ z^2 \\ \vdots \\ z^{n_t} \end{bmatrix} = H^m f^m = \begin{bmatrix} h^{m,1} \\ h^{m,2} \\ \vdots \\ h^{m,n_t} \end{bmatrix} f^m \quad (4.14)$$

### 4.3.3 Expectation Maximization

Expectation Maximization (EM) is a technique used to estimate a parameter  $X$  to maximize a joint probability density  $p(X, Z, K)$ , where  $Z$  is the data and  $K$  is termed as the missing data. In the case when jointly maximizing  $p(X, Z, K)$  with respect to  $X, K$  is not feasible, but maximizing the density with respect to  $X$  when  $K$  is known is possible, the EM technique can be used to iteratively estimate  $X$ .  $p(X, Z, K)$  is termed as the complete data likelihood and  $p(K|Z, \hat{X})$  is termed as the likelihood of missing data in this formulation. The EM technique proceeds as follows -

1.  $\hat{X}$  is a given initial estimate of  $X$
2. The missing data likelihood  $p(K|Z, \hat{X})$  is calculated using the previous estimate
3. The auxiliary function  $\mathcal{L}(X|\hat{X})$  is calculated as

$$\mathcal{L}(X|\hat{X}) = E_K \left\{ \log \left( p(X, Z, K) \right) \right\} \quad (4.15)$$

The expectation is with respect to the missing data likelihood  $p(K|Z, \hat{X})$

4. The new estimate  $\hat{X}$  is calculated by maximizing the auxiliary function
5. Steps 2 to 4 are repeated until the change in the auxiliary function falls below a threshold

6. Steps 2 and 3 involve an expectation operation and maximization operation, respectively, and thus the technique is termed as Expectation Maximization

## 4.4 Derivation of the ET-PMHT equations

The PMHT equations for a point multitarget tracking scenario are derived in [17]. The derivation in the current work follows a similar procedure to derive the equations for extended target tracking with the shape described by face triangulations.

### 4.4.1 ET-PMHT for initialization with clutter

Let the measurements for initialization be  $Z = \{z^1, z^2, \dots, z^{n_0}\}$  and the faces to be estimated be  $F = \{f^1, f^2, \dots, f^m\}$ . The missing data are the unknown associations between the faces and the measurements represented by  $K = \{k^1, k^2, \dots, k^{n_0}\}$ , where  $k^i$  takes values from 1 to M and denotes that the measurement  $z^i$  is associated to face  $f^{k^i}$ . The faces can be represented in terms of vector of vertices  $X_S$  using a selection matrix  $F^{m \times x}$  such that  $f^m = F^{m \times x} X_S$ . Let  $P_C$  the probability of clutter. The complete data likelihood  $p(X_S, K, Z)$  and the conditional probability of the missing data  $p(K|\hat{X}_S, Z)$  are given by-

$$p(X_S, K, Z) = \prod_{r=1}^{n_0} \mathcal{N}(z^r; h^{k^r} f^{k^r}, R) \quad (4.16)$$

$$p(K|\hat{X}_S, Z) = \prod_{r=1}^{n_0} p(k^r|\hat{X}_S, z^r) \quad (4.17)$$

$$w^{m,r} = p(k^r = m|\hat{X}_S, z^r) = \frac{\mathcal{N}(z^r; h^m f^m, R)}{\sum_{l=1}^M \mathcal{N}(z^r; h^l f^l, R) + P_C} \quad (4.18)$$

The auxiliary function can be written as

$$\begin{aligned}\mathcal{L}(X_S|\hat{X}_S) &= E_K \left( \log(p(X_S, K, Z)) \right) \\ &= -\frac{1}{2} \sum_{m=1}^M (Z^m - H^m F^{mx} X_S)^T (R^m)^{-1} (Z^m - H^m F^{mx} X_S)\end{aligned}\quad (4.19)$$

$Z^m, H^m, R^m$  are the set of measurements, measurement function and the measurement noise covariance, respectively for the measurements with weights greater than a threshold  $\Gamma_{in}$ , i.e.  $Z^m = \{z^r | w^{m,r} \geq \Gamma_{in}\}$ .

The linear system obtained is then -

$$\left[ \sum_{m=1}^M (H^m F^{mx})^T (R^m)^{-1} (H^m F^{mx}) \right] X_S = \sum_{m=1}^M (H^m F^{mx})^T (R^m)^{-1} Z^m \quad (4.20)$$

The procedure is followed until the cost between the iterations  $l$  and  $l + 1$  falls below a certain threshold  $\tau_I$  or the maximum number of iterations is reached. The estimate with the minimum cost is chosen at the end of the iterations.

$$\left| \mathcal{L}(X_S|\hat{X}_S)|_{l+1} - \mathcal{L}(X_S|\hat{X}_S)|_l \right| < \tau_I \quad (4.21)$$

#### 4.4.2 State extension

It was seen from simulations that using a single parameter  $\Omega$  in the estimate for capturing the target rigid body rotation and the coordinated turn in the translation kinematics was not feasible using first order linearization. If the target rotation is significant but its coordinated turn translation is not significant with respect to

the target extent, the same parameter cannot capture both rotation and translation. Thus, the state was extended with separate parameters  $\Omega_{CT}$  for translation coordinated turn parameter and  $\Omega_R$  for target rigid body rotation parameter. The total state is now written as

$$\mathbf{x} = \left[ X_S^T \quad v^T \quad \Omega_R \quad \Omega_{CT} \right]^T \quad (4.22)$$

Accordingly, the process noise covariance and the  $\Gamma_4^{Nx}$  matrix are modified as well.

$$Q = \begin{bmatrix} Q & 0 \\ 0 & q_\Omega \end{bmatrix} \quad \Gamma_5^{Nx} = \begin{bmatrix} \Gamma_4^{Nx} & 0 \\ 0 & T \end{bmatrix} \quad (4.23)$$

### 4.4.3 ET-PMHT for update with clutter

The complete data for the Expectation Maximization step consists of the total measurement set for times  $t = 1$  to  $t : \mathbb{Z} = \{Z_1, Z_2, \dots, Z_T\}$ , the states to be estimated  $\mathbb{X} = \{x(0), x(1), \dots, x(t)\}$  and the missing data of the association terms denoted by  $\mathbb{K} = \{K_1, K_2, \dots, K_T\}$ . The individual terms of the total measurement set  $\mathbb{Z}$  are the measurement sets for each time  $t : Z_t = \{z_t^1, z_t^2, \dots, z_t^{n_t}\}$  and the individual terms of  $\mathbb{X}$  are the states at each time. Each face  $f_t^m$  of the state can be written in terms of the state using a  $F^{mx}$ , such that  $f_t^m = F^{mx}x(t)$ . The individual terms of the association terms  $\mathbb{K}$  are  $K_t = \{k_t^1, k_t^2, \dots, k_t^{n_t}\}$  where  $k_t^r$  denotes that the measurement  $z_t^r$  is associated to the face  $f_t^{k_t^r}$ . The number of faces or the face triangulations are not changed during the EM procedure.

The complete data likelihood is given by

$$p_{comp}(\mathbb{X}, \mathbb{K}, \mathbb{Z}) = p(\mathbb{Z}|\mathbb{X}, \mathbb{K})p(\mathbb{K}|\mathbb{X})p(\mathbb{X}) \quad (4.24)$$

where  $p(\mathbb{Z}|\mathbb{X}, \mathbb{K})$  consists of the measurement likelihood terms

$$p(\mathbb{Z}|\mathbb{X}, \mathbb{K}) = \prod_{t=1}^T \prod_{r=1}^{n_t} \mathcal{N}(z_t^r; h_t^{k_t^r}(f_t^{k_t^r}), R) \quad (4.25)$$

$p(\mathbb{K}|\mathbb{X})$  denotes the probability of the current association of the measurements to the target faces. It can be approximated using the probability of the effective number of measurements associated to each face given the average number of measurements from the face. The probability of each face is calculated for the current association and only the faces with probability above a threshold are used in the Expectation Maximization. This face management step is described in Section 4.5.1.1.

$p(\mathbb{X})$  describes the evolution of the state according to the known dynamics.

$$p(\mathbb{X}) = p(x(0)) \prod_{t=1}^T p(x(t)|x(t-1)) \quad (4.26)$$

The missing data likelihood can be written as

$$p(\mathbb{K}|\mathbb{Z}, \hat{\mathbb{X}}) = \prod_{t=1}^T \prod_{r=1}^{n_t} p(k_t^r = m | z_t^r, \hat{x}(t)) = \prod_{t=1}^T \prod_{r=1}^{n_t} w_t^{m,r} \quad (4.27)$$

where,

$$w_t^{m,r} = p(k^r = m | \hat{X}_S(t), z_t^r) = \frac{\mathcal{N}(z_t^r; h_t^m f_t^m, R)}{\sum_{l=1}^M \mathcal{N}(z_t^r; h_t^l f_t^l, R) + P_C} \quad (4.28)$$

The auxiliary function is

$$\begin{aligned}
 \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}}) &= E_{\mathbb{K}}\left(\log(p_{comp}(\mathbb{X}, \mathbb{K}, \mathbb{Z}))\right) \\
 &= E_{\mathbb{K}}\left(\sum_{t=1}^T \sum_{r=1}^{n_t} \log\{\mathcal{N}(z_t^r; h_t^{k_t^r}(f_t^{k_t^r}), R)\} + \sum_{m=1}^M \left[\log\{p(f_0^m)\} + \sum_{t=1}^T \log\{p(f_t^m|f_{t-1}^m)\}\right]\right)
 \end{aligned} \tag{4.29}$$

Assuming the measurement function to be linear and writing the face  $f_t^m$  in terms of the state, we obtain  $h_t^{k_t^r}(f_t^{k_t^r}) = h_t^{m,r} F^{m,x} x(t)$ . The density of the initial estimate is assumed to be Gaussian  $p(F^{m,x} x(0)) = \mathcal{N}(F^{m,x} x(0); F^{m,x} \bar{x}(0), P^m(0))$  and using (4.3), the transition density is  $p(F^{m,x} x(t)|F^{m,x} x(t-1)) = \mathcal{N}(x(t); f_{t-1}(x(t-1)), (F^{m,x})^T Q^{-1} F^{m,x})$ . Setting initial and final values of  $t$  as  $T-1$  and  $T$ , the auxiliary function, dropping the terms independent of the state, can be simplified to

$$\begin{aligned}
 \mathcal{L}(\mathbb{X}|\hat{\mathbb{X}}) &= -\frac{1}{2} \sum_{m \in M_T^{vis}} (Z_T^m - H_T^m F^{m,x} x(T))^T (R_T^m)^{-1} (Z_T^m - H_T^m F^{m,x} x(T)) \\
 &\quad - \frac{1}{2} \sum_{m=1}^M (x(T-1) - \bar{x}(T-1))^T (F^{m,x})^T (P^m(T-1))^{-1} F^{m,x} (x(T-1) - \bar{x}(T-1)) \\
 &\quad - \frac{1}{2} \sum_{m=1}^M (x(T) - f_{T-1}(x(T-1)))^T (F^{m,x})^T Q^{-1} F^{m,x} (x(T) - f_{T-1}(x(T-1)))
 \end{aligned} \tag{4.30}$$

where  $Z_T^m, H_T^m$  and  $R_T^m$  are the stacked vectors/matrix are defined as

$$Z_T^m = \begin{bmatrix} z_T^1 \\ z_T^2 \\ \vdots \\ z_T^{n_m} \end{bmatrix}, H_T^m = \begin{bmatrix} h_T^{m,1} \\ h_T^{m,2} \\ \vdots \\ h_T^{m,n_m} \end{bmatrix}, (R_T^m)^{-1} = \begin{bmatrix} w_T^{m,1} R^{-1} & & & \\ & w_T^{m,2} R^{-1} & & \\ & & \ddots & \\ & & & w_T^{m,n_m} R^{-1} \end{bmatrix} \quad (4.31)$$

$M_T^{vis}$  is the set of the visible faces at time  $T$ .  $Z_T^m$  are the measurements with weights  $w_T^{m,r}$  greater than  $\Gamma_{in}$  and  $H_T^m, R_T^m$  are the corresponding measurement functions and measurement noise covariances.

The transition function (4.2) can be simplified as

$$x(t+1) = \begin{bmatrix} [F_R^{N_x}(I - V^{N_x} M^{N_x}) + V^{N_x} M^{N_x}](X_S(t) + V^{N_x} F_{pv}v(t)) \\ F_{vv}v(t) \end{bmatrix} \quad (4.32)$$

$$\text{with } V^{N_x} = \begin{bmatrix} I_3 \\ I_3 \\ \vdots \\ I_3 \end{bmatrix} \text{ and } M^{N_x} = \frac{1}{N_x} \begin{bmatrix} I_3 & I_3 & \dots & I_3 \end{bmatrix} \quad (4.33)$$

$F_R^{N_x}$  is a function of  $\Omega_R$  and  $F_{pv}$  and  $F_{vv}$  are functions of  $\Omega_{CT}$ . The transition function can be further simplified by noting that  $M^{N_x} V^{N_x} = I_3$ .

$$x(t+1) = \begin{bmatrix} [F_R^{N_x}(I - V^{N_x} M^{N_x}) + V^{N_x} M^{N_x}](X_S(t)) + V^{N_x} F_{pv}v(t) \\ F_{vv}v(t) \end{bmatrix} = f_t(x(t)) \quad (4.34)$$

Linearization about the previous estimate  $\hat{x}(t)$  is used to handle the nonlinear transition function  $f_t(x(t))$ .

$$f_t(x(t)) = f_t(\hat{x}(t)) + J(x(t) - \hat{x}(t)), \quad (4.35)$$

$$\text{where } J \text{ is the Jacobian defined as } J = \left. \frac{\partial f_t(x(t))}{\partial x(t)} \right|_{x(t)=\hat{x}(t)}$$

The linearized auxiliary function can be obtained by replacing the nonlinear transition function with the linear version given in (4.36). The optimal value to maximize the linearized auxiliary function is obtained by taking the derivative of the auxiliary function with respect to  $x_{T-1}$ ,  $x_T$  and equating it to zero. The final linear system obtained is -

$$\begin{bmatrix} \sum_{m=1}^M (F^{m,x})^T (P^m(T-1))^{-1} F^{m,x} + J^T Q_m^{-1} J & -\sum_{m=1}^M J^T Q_m^{-1} \\ -\sum_{m=1}^M Q_m^{-1} J & \sum_{m \in M_T^{vis}} (H_T^m F^{m,x})^T (R_T^m)^{-1} (H_T^m F^{m,x}) + \sum_{m=1}^M Q_m^{-1} \end{bmatrix} \begin{bmatrix} x(T-1) \\ x(T) \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix} \quad (4.36)$$

where

$$Q_m^{-1} = (F^{m,x})^T Q^{-1} F^{m,x} \quad (4.37)$$

$$\begin{aligned} A &= \sum_{m=1}^M \left[ (F^{m,x})^T (P^m(T-1))^{-1} F^{m,x} \right] \bar{x}(T-1) \\ &\quad + \sum_{m=1}^M (-J^T Q_m^{-1}) (-f_{T-1}(\hat{x}(T-1)) + J\hat{x}(T-1)) \end{aligned} \quad (4.38)$$

$$B = \sum_{m \in M_T^{vis}} \left[ (H_T^m F^{m,x})^T (R_T^m)^{-1} \right] Z_T^m + \sum_{m=1}^M (Q_m^{-1}) (-f_{T-1}(\hat{x}(T-1)) + J\hat{x}(T-1)) \quad (4.39)$$

The derivation of the Jacobian of the transition function  $J$  is given in Appendix 4.8. It was seen from the simulations that implementing the total linear system directly is not stable and gives erroneous results. The problem is split into two sub-problems - the kinematics are estimated while keeping the shape fixed in Section 4.4.4 and updating the only the shape using the current measurements in Section 4.4.5.

#### 4.4.4 ET-PMHT with fixed shape

In a similar manner, ET-PMHT equations can be obtained keeping the shape fixed and updating only the kinematics using the measurements. The update is split into measurement update and transition update, and the ET-PMHT is used only for the update using the measurements. It can be seen from Equation (4.34) that the updated shape is only a function of  $v(T - 1)$ . The complete data likelihood is then given by-

$$p(v(T - 1), K_T, Z_T) = p(Z_T|K_T, v(T - 1))p(K_T|v(T - 1))p(v(T - 1)) \quad (4.40)$$

$$p(Z_T|K_T, v(T - 1)) = \prod_{r=1}^{n_t} \mathcal{N}(z_T^r; h_T^{k_T^r, r} F^{m_x} X_S(T), R) \quad (4.41)$$

$$p(v(T - 1)) = \mathcal{N}(v(T - 1), \bar{v}(T - 1), P_v(T - 1)) \quad (4.42)$$

$p(K_T|v(T - 1))$  is approximated using the probability of actual number of measurements obtained from the target versus the expected number of measurements from the target when the shape is updated using  $v(T - 1)$ , and handled in a similar

manner as before. The auxiliary function is then-

$$\begin{aligned}
 \mathcal{L}(v(T-1)|\hat{v}(T-1)) &= \\
 &- \frac{1}{2} \sum_{m \in M^{vis}} (Z_T^m - H_T^m F^{mx} X_S(T))^T (R_T^m)^{-1} (Z_T^m - H_T^m F^{mx} X_S(T)) \\
 &- \frac{1}{2} (v(T-1) - \bar{v}(T-1))^T (P_v(T-1))^{-1} (v(T-1) - \bar{v}(T-1)) \quad (4.43)
 \end{aligned}$$

The final linear system obtained is

$$\begin{aligned}
 &\left[ (P_v(T))^{-1} + \sum_{m \in M^{vis}} (H_T^m F^{mx} J_{xv})^T (R_T^m)^{-1} (H_T^m F^{mx} J_{xv}) \right] v(T-1) = \\
 &(P_v(T))^{-1} \bar{v}(T-1) \\
 &\quad + \sum_{m \in M^{vis}} (H_T^m F^{mx} J_{xv})^T (R_T^m)^{-1} \left( Z_T^m - (f_x(\hat{v}(T-1)) - J_{xv} \hat{v}(T-1)) \right) \quad (4.44)
 \end{aligned}$$

with  $X_S(T) = f_x(v(T-1)), J_{xv} = \frac{\partial X_S(T-1)}{\partial v(T-1)} = \left[ V^{N_x} F_{pv} \quad \nabla_{\Omega_{CT}} X_S(T) \quad \nabla_{\Omega_R} X_S(T) \right]$ .

The procedure is followed until the cost between the iterations  $l$  and  $l+1$  falls below a certain threshold  $\tau_{FS}$  or the maximum number of iterations are reached.

$$\left| \mathcal{L}(v(T-1)|\hat{v}(T-1))|_{l+1} - \mathcal{L}(v(T-1)|\hat{v}(T-1))|_l \right| < \tau_{FS} \quad (4.45)$$

The transition update is performed as follows using the parameters obtained from the iteration with the minimum cost.

$$v(T) = f_v(v(T-1)) \quad (4.46)$$

$$P(T) = JP(T)J^T + \Gamma_5^{N_x} Q(\Gamma_5^{N_x})^T \quad (4.47)$$

#### 4.4.5 ET-PMHT shape update

Using the predicted shape using the updated kinematics from Section 4.4.4, as the initial shape estimate, the auxiliary function for the shape update is written as

$$\begin{aligned}
 \mathcal{L}(X_S(T)|\hat{X}_S(T)) = & \\
 & - \frac{1}{2} \sum_{m=1}^M (X_S(T) - \bar{X}_S(T))^T (F^{mx})^T (P_S^m(T))^{-1} F^{mx} (X_S(T) - \bar{X}_S(T)) \\
 & - \frac{1}{2} \sum_{m \in M_T^{vis}} (Z_T^m - H_T^m F^{mx} X_S(T))^T (R_T^m)^{-1} (Z_T^m - H_T^m F^{mx} X_S(T)) \quad (4.48)
 \end{aligned}$$

In a similar manner as in the previous sections, the linear system can be derived as -

$$\begin{aligned}
 \left[ \sum_{m=1}^M (F^{mx})^T (P_S^m(T))^{-1} F^{mx} + \sum_{m \in M_T^{vis}} (H_T^m F^{mx})^T (R_T^m)^{-1} (H_T^m F^{mx}) \right] X_S(T) = \\
 \sum_{m=1}^M (F^{mx})^T (P_S^m(T))^{-1} F^{mx} \bar{X}_S(T) + \sum_{m \in M_T^{vis}} (H_T^m F^{mx})^T (R_T^m)^{-1} Z_T^m \quad (4.49)
 \end{aligned}$$

The procedure is followed until the cost between the iterations  $l$  and  $l + 1$  falls below a certain threshold  $\tau_{SU}$  or the maximum number of iterations are reached.

$$\left| \mathcal{L}(X_S(T)|\hat{X}_S(T))|_{l+1} - \mathcal{L}(X_S(T)|\hat{X}_S(T))|_l \right| < \tau_{SU} \quad (4.50)$$

The covariance of the shape is updated as follows using the parameters from the iteration with minimum cost.

$$S(T) = \mathcal{H}_T P_S(T) \mathcal{H}_T^T + \mathcal{R}_T \quad (4.51)$$

$$W(T) = P_S(T) \mathcal{H}_T^T S(T)^{-1} \quad (4.52)$$

$$P_S(T) = P_S(T) - W(T) S(T) W(T)^{-1} \quad (4.53)$$

where  $\mathcal{H}_T$  is the vertically stacked matrix with elements  $H_T^m F^{mx}$  and  $\mathcal{R}_T$  is the block diagonal matrix with elements  $R_T^m$  for visible faces  $m$ . The total covariance is then given by

$$P(T) = \begin{bmatrix} P_S(T) & \\ & P_v(T-1) \end{bmatrix} \quad (4.54)$$

It must be noted that the covariance for the kinematics is unchanged. This is to capture the coordinated turn dynamics by inflating the covariance as described in [18].

#### 4.4.6 Face to Face association

An association step is required to maintain the association between the faces of the estimate across time steps. Since the number of faces may also change across time steps, a many-to-one association is performed, i.e. multiple faces at time step  $t-1$  can be associated to a single face at time step  $t$ . A distance metric is required to measure the similarity between the face triangulations. The Hausdorff distance is chosen as the distance metric.

#### **4.4.6.1 Hausdorff Distance**

The Hausdorff distance is a metric to measure the similarity between 2 convex sets A and B with a distance measure  $d(x, y)$  defined between the elements  $x \in A$  and  $y \in B$ . It is defined as

$$d_{Hausdorff}(A, B) = \max_{x \in A}(\min_{y \in B}(d(x, y))) \quad (4.55)$$

The Hausdorff distance between 2 triangles A and B can thus be calculated as the maximum distance between the vertices of triangle A to the vertices of triangle B.

## **4.5 Algorithm**

### **4.5.1 Initialization**

#### **4.5.1.1 Shape Initialization**

Given the measurements  $Z_0$  at the initial time frame 0, the objective is to find a shape estimate  $\hat{X}_S(0)$ . It must be noted that the measurements include clutter as well, and the algorithm cannot distinguish a priori whether a measurement is from the target or from clutter. The initial shape estimate  $\hat{X}_S^0(0)$  for the algorithm is calculated as follows-

- The shape estimate is initially calculated as the convex hull of the entire set of measurements  $Z_0$ .

- Face management : The quality  $q_m$  of each face  $m$  of the estimate is calculated as -

$$q_m = Poisson(n_m, n_{avg}) \quad (4.56)$$

$$n_m = \sum_{r=1}^{n_0} w_0^{m,r} \quad (4.57)$$

The faces with quality lesser than a threshold  $\Gamma_I$  are deleted from the shape estimate to obtain the initial shape estimate  $\hat{X}_S^0(0)$ .

The following operations are performed until the cost between two iterations  $l$  and  $l + 1$  falls below a tolerance limit or until the maximum number of iterations is reached.

- ET-PMHT for initialization (Section 4.4.1) is used to obtain a filtered estimate  $X_S^l(0)$  for the  $l^{th}$  iteration.
- Face management is used to delete the erroneous faces after the filtering.
- The convex hull operation is performed on the estimate and only the visible faces are selected.
- The cost for the visible faces is calculated according to Equation (4.21).
- The measurements are validated using only the visible faces and only the measurements falling within the validation region are selected for initializing new faces. The measurements are validated, using their respective weights

from Equation (4.18), as follows -

$$\Gamma_{out} \leq w^{m,r} < \Gamma_{in} \quad (4.58)$$

- Initializing new faces - The validated measurements are grouped into sets according to the edge of the visible faces they are closest to. Each set is used to initialize a new face with the vertices being the endpoints of the closest edge and the farthest measurement in the set. Face management is used again to delete the faces that may have been erroneously added.

The final estimate  $\hat{X}_S(0), \hat{X}_{DT}(0)$  at the end of the iterations is the one with the minimum cost. The initialization pseudocode is given below in Algorithm 5.

#### 4.5.1.2 State Initialization

The translational kinematics  $v_x, v_y, v_z$  are initialized as the difference in the mean of the measurements from consecutive time steps and the rotational parameters  $\Omega_R$  and  $\Omega_{CT}$  are set to 0. The total state estimate and its covariance is given by -

$$\hat{\mathbf{x}}(0) = \left[ \hat{X}_S^T(0) \quad \hat{v}_x(0) \quad \hat{v}_y(0) \quad \hat{v}_z(0) \quad 0 \quad 0 \right]^T \quad P(0) = \begin{bmatrix} P_S(0) & \\ & P_v(0) \end{bmatrix} \quad (4.59)$$

Additional faces are added to the estimate during the convex hull operation. To differentiate between the faces that are actually estimated using measurements and those that are additional faces added due to convex hull, an indicator variable  $\delta_O$  is used.  $\delta_O = 1$  implies that the face is valid and has been estimated and  $\delta_O = 0$

---

**Algorithm 5** Shape Initialization

---

```

procedure INITIALIZE_SHAPE( $Z(0), \Gamma_I, \Gamma_G, \text{iter\_max}$ )
     $X_S^0(0), X_{DT(0)}^0 \leftarrow \text{get\_convex\_hull}(Z(0))$      $L_X, L_{DT}, L_C \leftarrow [], [], []$ 
     $X_S^0(0), X_{DT(0)}^0 \leftarrow \text{face\_management}(\hat{X}_S^0(0), \hat{X}_{DT}^0(0), Z(0), \Gamma_I)$ 
    for  $l=1, l < \text{iter\_max}$  do
         $X_S^l(0), X_{DT(0)}^l \leftarrow \text{ET\_PMHT\_initialization}(\hat{X}_S^{l-1}(0), \hat{X}_{DT}^{l-1}(0), Z(0))$ 
         $X_S^l(0), X_{DT(0)}^l \leftarrow \text{face\_management}(\hat{X}_S^l(0), \hat{X}_{DT}^l(0), Z(0), \Gamma_I)$ 
         $X_S^l(0), X_{DT(0)}^l \leftarrow \text{get\_convex\_hull}(\hat{X}_S^l(0))$ 
         $X_S^l(0), X_{DT(0)}^l \leftarrow \text{get\_visible\_faces}(\hat{X}_S^l(0), \hat{X}_{DT}^l(0))$ 
         $C^l \leftarrow \text{init\_cost}(\hat{X}_S^l(0), \hat{X}_{DT}^l(0), Z(0))$ 
         $L_X, L_{DT}, L_C \leftarrow [L_X \ X_S^l(0)], [L_{DT} \ X_{DT(0)}^l], [L_C \ C^l]$ 
         $Z_G(0) \leftarrow \text{validate\_measurements}(\hat{X}_S^l(0), \hat{X}_{DT}^l(0), Z(0))$ 
         $X_S^l(0), X_{DT(0)}^l \leftarrow \text{initialize\_new\_faces}(\hat{X}_S^l(0), \hat{X}_{DT}^l(0), Z_G(0))$ 
         $X_S^l(0), X_{DT(0)}^l \leftarrow \text{face\_management}(\hat{X}_S^l(0), \hat{X}_{DT}^l(0), Z(0), \Gamma_I)$ 
     $\hat{X}_S(0), \hat{X}_{DT(0)} \leftarrow \text{min\_cost\_estimate}(L_X, L_{DT}, L_C)$ 

```

---

implies that the face is an extension to form the convex hull. An array of these indicator variables represented by  $\Delta_O(0)$  is initialized to 1 for each of its elements.

## 4.5.2 Update

Given the estimates from frame  $t$ ,  $\mathbf{x}(t)$ ,  $P(t)$ ,  $X_{DT}(t)$ ,  $\Delta_O(t)$ , and measurements  $Z(t+1)$  at frame  $t+1$ , the objective is to find the updated estimates  $\mathbf{x}(t+1)$ ,  $P(t+1)$ ,  $X_{DT}(t+1)$ ,  $\Delta_O(t+1)$ .

- ET-PMHT with fixed shape described in Section 4.4.4 is used to obtain the updated state  $\mathbf{x}_{FS}(t+1)$  and covariance  $P_{FS}(t+1)$  with the shape unchanged from the previous frame. The cost for this estimate is  $C_{FS}$ .
- Using  $\mathbf{x}_{FS}(t+1)$  as the initial estimate and  $P_{FS}(t+1)$  as the covariance, ET-PMHT shape update described in Section 4.4.5 is used to obtain the state with updated shape  $\mathbf{x}_{SU}(T)$ , updated covariance  $P_{SU}(T)$  and cost  $C_{SU}$ .
- The estimate with the lower cost among  $C_{FS}$  and  $C_{SU}$  is chosen with the appropriate covariance.
- The measurements are validated using Equation (4.58) with weights obtained from Equation (4.28) and the measurements are chosen to initialize new faces.
- The measurements chosen are clustered using DBSCAN to obtain clusters while handling clutter. A plane is estimated for each cluster using the measurements in the cluster. New faces are initialized for each cluster as the 2D convex hull of the projection of the measurements onto the plane estimated.

- Face Management is used to delete faces with low quality, Equation (4.56), as described in Section 4.5.1.1.
- The convex hull operation is performed to constrain the updated estimate to a convex shape.  $\Delta_O(t+1)$  is obtained by associating the faces with  $\delta_O = 1$  in the updated shape to the faces in the convex hull.

---

**Algorithm 6** Update

---

```

procedure UPDATE( $Z(t+1), \hat{\mathbf{x}}(t), P(t), \Delta_O(t), \Gamma_G$ )
     $\hat{\mathbf{x}}_{FS}(t+1), P_{FS}(t+1), C_{FS} \leftarrow$  ET-PMHT-FS( $\hat{\mathbf{x}}(t), P(t), \Delta_O(t)$ )
     $\hat{\mathbf{x}}_{SU}(t+1), P_{SU}(t+1), C_{SU} \leftarrow$  ET-PMHT-SU( $\hat{\mathbf{x}}_{FS}(t+1), P_{FS}(t+1), \Delta_O(t)$ )
     $\hat{\mathbf{x}}(t+1), P(t+1) \leftarrow$  min_cost_estimate( $\hat{\mathbf{x}}_{FS}(t+1), P_{FS}(t+1), C_{FS}, \hat{\mathbf{x}}_{SU}(t+1), P_{SU}(t+1), C_{SU}$ )
     $Z_I \leftarrow$  validate_measurements( $\hat{\mathbf{x}}(t+1), P(t+1), \Delta_O(t), \Gamma_G$ )
     $\hat{\mathbf{x}}(t+1), P(t+1), \Delta_O(t+1) \leftarrow$  initialize_new_faces( $Z_I$ )
     $\hat{\mathbf{x}}(t+1), P(t+1), \Delta_O(t+1) \leftarrow$  face_management( $\hat{\mathbf{x}}(t+1), P(t+1), \Delta_O(t+1)$ )
     $\hat{\mathbf{x}}(t+1), P(t+1), \Delta_O(t+1) \leftarrow$  convex_hull( $\hat{\mathbf{x}}(t+1), P(t+1), \Delta_O(t+1)$ )

```

---

## 4.6 Results

The performance of the algorithm is tested for different scenarios and the following metrics are used to quantify the performance

- RMSE of the center - The center of the target and the estimate for the time frame  $t$  are calculated as the mean of their vertices and represented as  $c_T(t)$

and  $c_E(t)$  respectively. The RMSE of the center denoted by  $e_C(t)$  is then

$$e_C(t) = \sqrt{|c_T(t) - c_E(t)|^2}$$

- RMSE of the kinematics - The RMSE of the kinematics denoted by  $e_K(t)$  is calculated in a similar manner with  $v_T(t), \Omega(t)$  as the target kinematics and  $v_E(t), \Omega_R(t), \Omega_{CT}(t)$  as the estimate kinematics.

$$e_K(t) = \sqrt{\left| \left[ \begin{array}{ccc} (v_T(t))^T & \Omega(t) & \Omega(t) \end{array} \right] - \left[ \begin{array}{ccc} (v_E(t))^T & \Omega_R(t) & \Omega_{CT}(t) \end{array} \right] \right|^2}$$

- Intersection over Union (IoU) - The Intersection over Union is a metric used to quantify the similarity between 2 shapes. The metric is defined for shapes  $S_1$  and  $S_2$  as

$$\text{Intersection over Union} = \frac{V(S_1 \cap S_2)}{V(S_1 \cup S_2)}$$

where,  $V(S_1 \cap S_2)$  is the volume of the intersection of shapes and  $V(S_1 \cup S_2)$  is the volume of the union of the shapes. In the current work, the denominator is approximated by the convex hull of the union of the shapes, since the union of the convex shapes may not be convex. It must be noted that this approximation is pessimistic and becomes closer to the true value as the estimate shape is closer to the target shape.

The performance of the proposed algorithm is compared with the 3D Gaussian Process, (Section 4.6.1) under the following scenarios -

- Different levels of measurement noise covariance
- Different levels of average number of measurements per unit area
- Different levels of clutter density

### 4.6.1 3D Gaussian Process

The extent of the target in this approach is described using a radial function which follows a Gaussian Process

$$f(\theta, \phi) \sim \mathcal{GP}(\mu(\theta, \phi), \kappa(\gamma, \gamma')) \quad (4.60)$$

where  $\gamma = (\theta, \phi)$ . The covariance function  $\kappa(\gamma, \gamma')$  describes the relation between different parts of the target given by

$$\kappa(\gamma, \gamma') = \sigma_f^2 e^{-\frac{d^2(\gamma, \gamma')}{2l^2}} + \sigma_r^2 \quad (4.61)$$

$$d(\gamma, \gamma') = \cos^{-1} \left( \cos(\phi) \cos(\phi') \cos(\theta) \cos(\theta') + \cos(\phi) \cos(\phi') \sin(\theta) \sin(\theta') \right. \\ \left. + \sin(\phi) \sin(\phi') \right) \quad (4.62)$$

where  $l$  is the length scale,  $\sigma_f^2$  is the prior variance and  $d(\gamma, \gamma')$  is the distance function between two angle pairs. The process model is given by

$$\mathbf{x}(t+1) = F(t)\mathbf{x}(t) + \nu(t) \sim \mathcal{N}(0, Q) \quad \mathbf{x}(t) = \begin{bmatrix} x_t(t) & x_r(t) & f(t) \end{bmatrix} \quad (4.63)$$

$x_t(t)$  denotes the translation component of the state with  $c(t)$  as the center,  $x_r(t)$  denotes the rotational component of the state and  $f(t)$  encodes the shape as a radial function. The measurement model is given by

$$z_{tl} = \tilde{h}(\mathbf{x}(t), z_{tl}) + e_{tl} \sim \mathcal{N}(0, R_{tl}) \quad (4.64)$$

$$\tilde{h}(\mathbf{x}(t), z_{tl}) = c(t) + p_{tl} H^f(\gamma_{tl}(c(t), q_t, z_{tl})) \quad (4.65)$$

$$H^f(\gamma_{tl}) = \kappa(\gamma_{tl}, \gamma_f) \kappa(\gamma_f, \gamma_f)^{-1} \quad (4.66)$$

$$R_{tl} = p_{tl} R_{tl}^f p_{tl}^T + R \quad (4.67)$$

Using initial estimate as  $\hat{\mathbf{x}}(0) = \mathcal{N}(\bar{\mathbf{x}}(0), P(0))$ , an EKF is used to update the state.

## 4.6.2 Scenario

The target shape is shown in Figure 4.2 and a sample trajectory of the scenario is shown in Figure 4.3. The target moves in such a manner that different faces of the target are visible to the sensor at different times. The sampling time between frames  $T = 1s$ . The sensor is located at  $\begin{bmatrix} 0m & 0m & 0m \end{bmatrix}^T$  and the target is centered at  $\begin{bmatrix} -14m & 10m & 0.2m \end{bmatrix}^T$ , with initial velocity  $\begin{bmatrix} 1.5ms^{-1} & 0 & -0.1ms^{-1} \end{bmatrix}^T$ . The target performs two maneuvers with  $\Omega = \frac{\pi}{100} rad s^{-1}$  from frames 5 to 35 and with  $\Omega = -\frac{\pi}{50} rad s^{-1}$  from frames 40 to 45. 20 Monte Carlo runs are used to average the performance metrics.

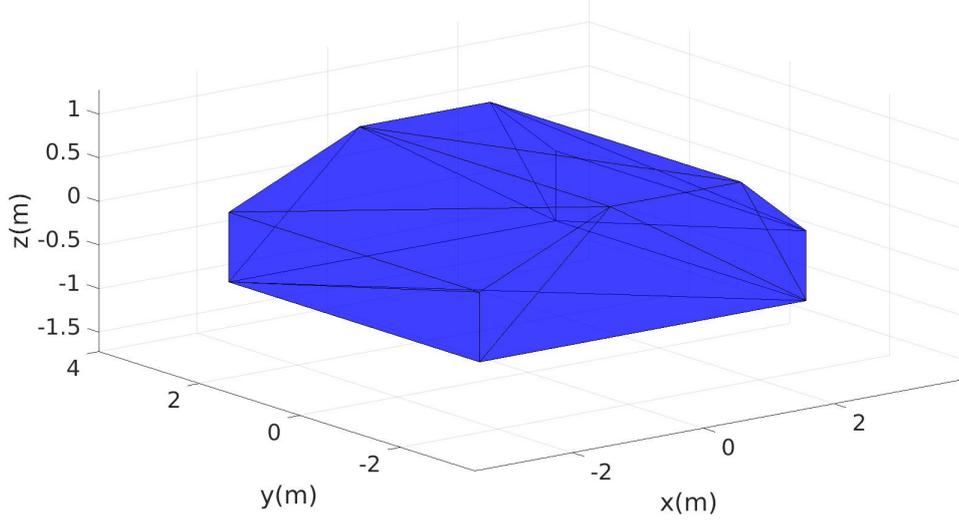


FIGURE 4.2: Target Shape

The measurements noise covariance and power spectral density of the process noise covariance are given below

$$R = \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{bmatrix}, q_x = q_y = q_z = 10^{-4} m^2 s^{-3}, q_\Omega = 10^{-6} rad^2 s^{-3}$$

The center of the cuboid in which the clutter is generated is located at the target center for each frame. The length, width and the height of the cuboid are  $7m, 7m, 2m$  respectively. The cuboid is divided into bins in which the clutter points can occur, with  $C_x = C_y = 100$  and  $C_z = 10$ .

The tracker parameters used are -

- Threshold for gating -  $\Gamma_{in} = 0.1 |M_t^{vis}|^{-1}$  and  $\Gamma_{out} = 10^{-10}$

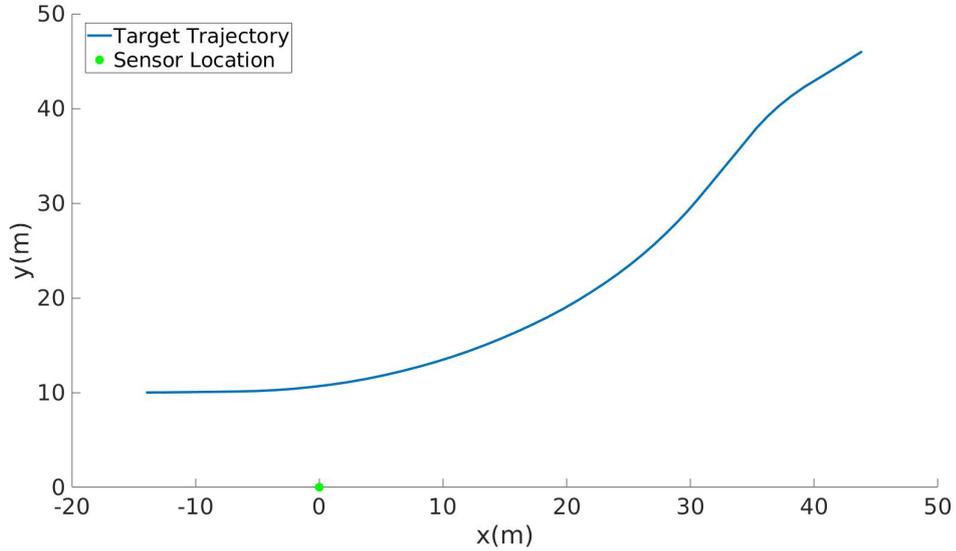


FIGURE 4.3: Sample Trajectory

- Parameters for DBSCAN -  $\epsilon = 1.2m$ , number of points = 5
- Threshold for Face Management -  $\tau_{FM} = 0.1$
- Tolerance for EM = 1
- Maximum number of iterations for EM = 10

The parameters used for the Gaussian Process are  $\sigma_r = 0.2m$ ,  $\sigma_f = 1m$ ,  $l = \frac{\pi}{8}$ , and 642 basis points are chosen for the shape function.

### 4.6.3 Different levels of measurement noise covariance

The algorithms are tested for various levels of measurement noise covariance while keeping the other parameters fixed. The density of measurements  $\rho = 3m^{-2}$  and

the clutter parameter  $\rho_C = 0.1m^{-3}$ . The standard deviation of the measurement noise covariance  $\sigma_x = \sigma_y = \sigma_z$  is varied among the values  $\{0.05m^{-1}, 0.1m^{-1}, 0.15m^{-1}\}$

Figures 4.4,4.5,4.6 show the RMSE of the center, RMSE of the kinematics and the IoU results. The snapshots of the ground truth, measurements and the estimates at different frames for a sample run is shown in Figure 4.7.

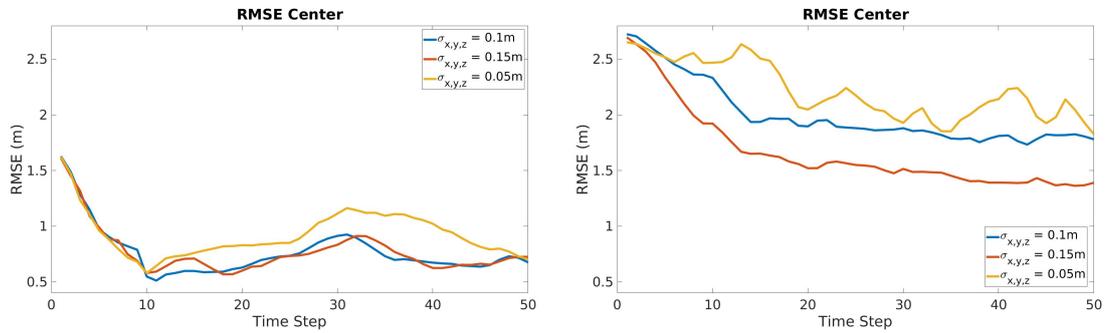


FIGURE 4.4: RMSE center for different levels of measurement noise covariance. Results for the proposed approach are on the left and the GP results are on the right

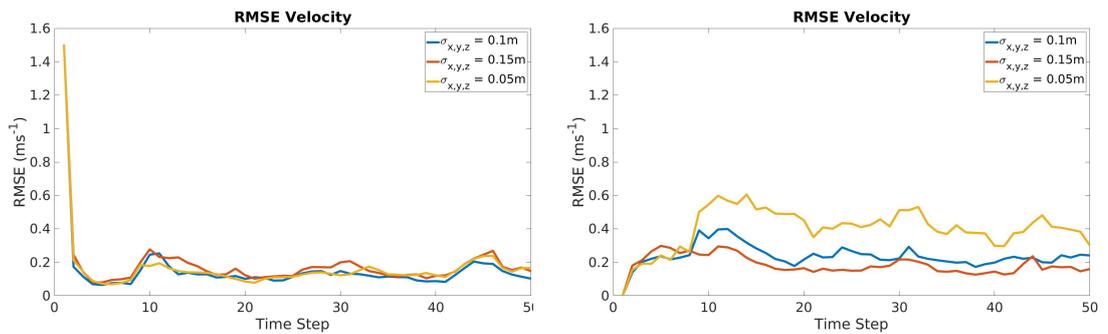


FIGURE 4.5: RMSE velocity for different levels of measurement noise covariance. Results for the proposed approach are on the left and the GP results are on the right

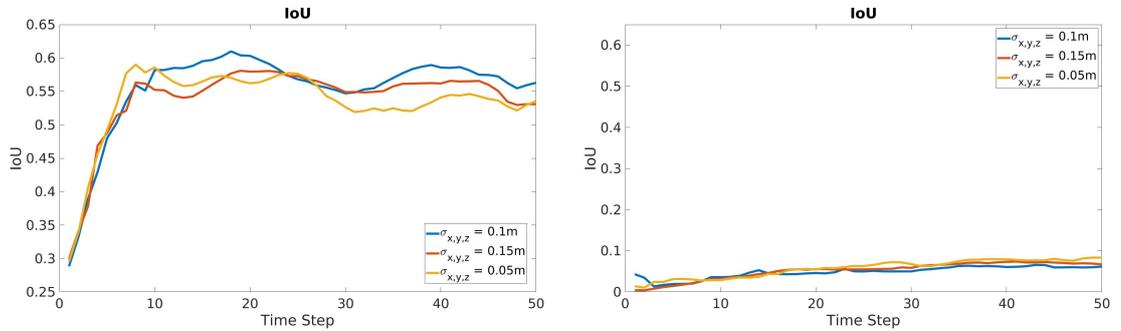


FIGURE 4.6: IoU results for different levels of measurement noise covariance. Results for the proposed approach are on the left and the GP results are on the right

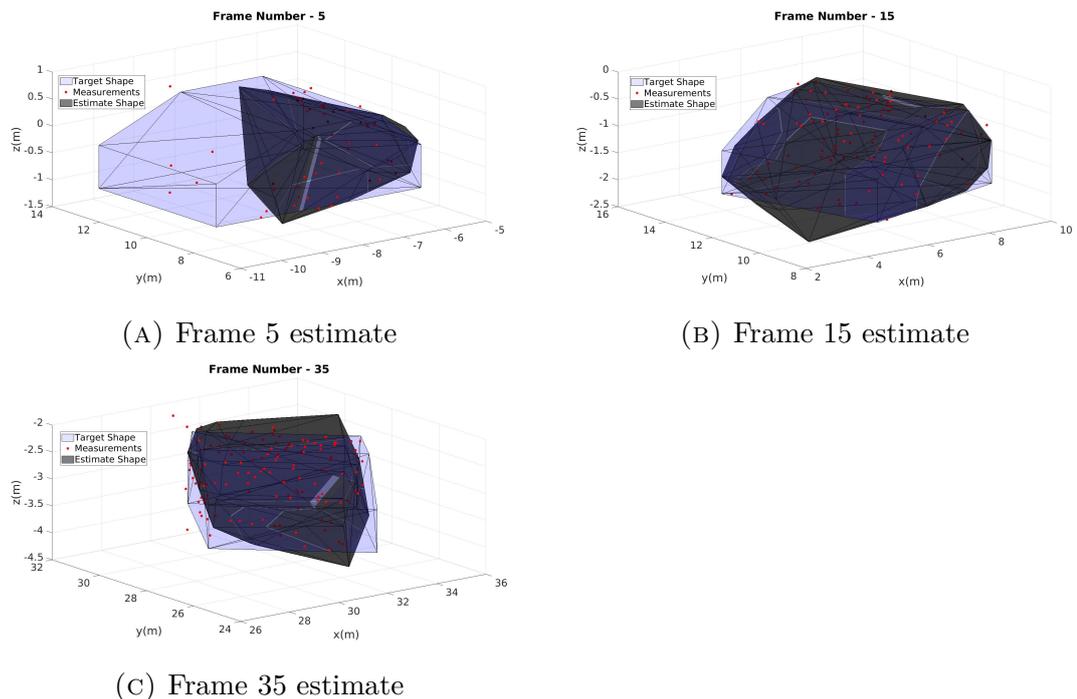


FIGURE 4.7: Estimates of the target shape at different times  $\sigma_x = \sigma_y = \sigma_z = 0.1m$ ,  $\rho = 3m^{-2}$  and  $\rho_C = 0.1m^{-3}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using proposed algorithm is in black.

The results show that the proposed algorithm is able to track the maneuvering target even with self-occlusion and in presence of clutter. It can be seen from Figure 4.7 that at frame 5, only a part of the target is estimated since only this part was visible until that frame. It can also be seen that the clutter is rejected and not used in the shape estimation. New faces are added as more parts of the target become visible as seen in frames 15 and 35. It can be seen that the estimate is able to track the visible parts more effectively than the parts that are not visible. The estimate shape is biased towards the parts of the target that generate no measurements since there are no constraints from the measurements to be applied to them. Yet another effect is when a face of the estimate is visible but the corresponding target face is not. No measurements will be generated and it may lead to deletion of a valid estimate face. As such, the effect of the shape uncertainty dominates and the effect of the change in the measurement noise covariance on the performance is not significant.

The Gaussian Process is unable to track the target especially due to self-occlusion. It can be seen from the results, particularly IoU, that the shape estimate of GP is severely affected by self-occlusion and the performance is significantly degraded.

#### **4.6.4 Different levels of average number of measurements per unit area**

The average number of measurements per unit area  $\rho$  takes values in the set  $\{1m^{-2}, 2m^{-2}, 3m^{-2}\}$ , while the other parameters are fixed,  $\sigma_x = \sigma_y = \sigma_z = 0.1m$ ,

and  $\rho_C = 0.1m^{-3}$ . Figures 4.8,4.9,4.10 show the RMSE of the center, RMSE of the kinematics and the IoU results. The snapshots of the ground truth, measurements and the estimates at different frames for a sample run is shown in Figure 4.11.

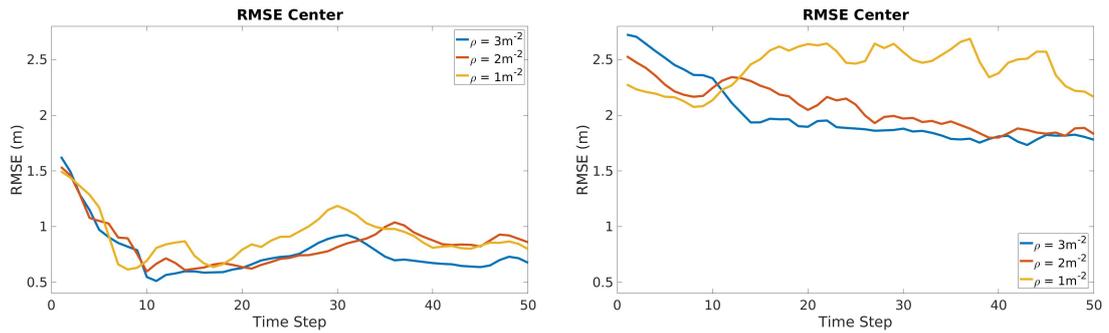


FIGURE 4.8: RMSE center for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right

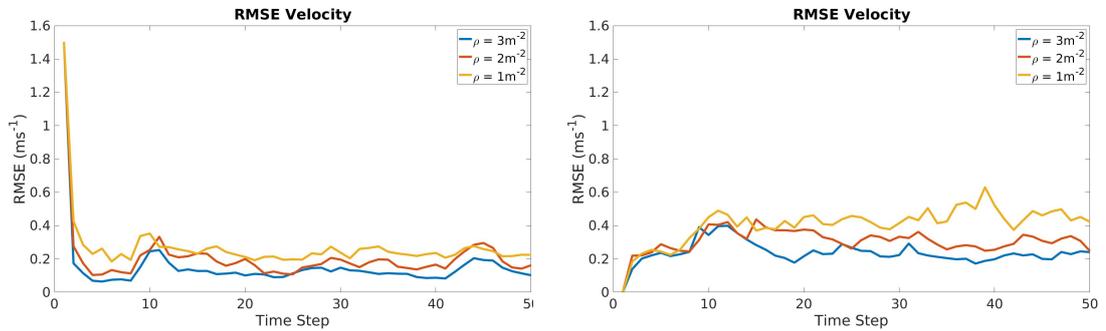


FIGURE 4.9: RMSE velocity for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right

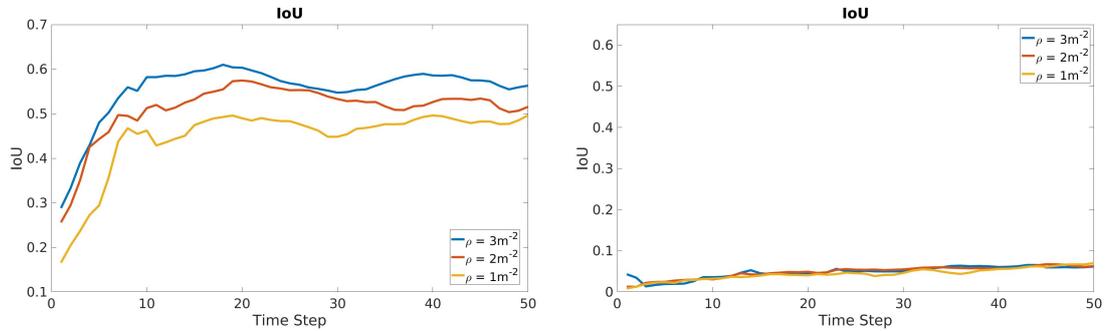


FIGURE 4.10: IoU results for different levels of measurements. Results for the proposed approach are on the left and the GP results are on the right

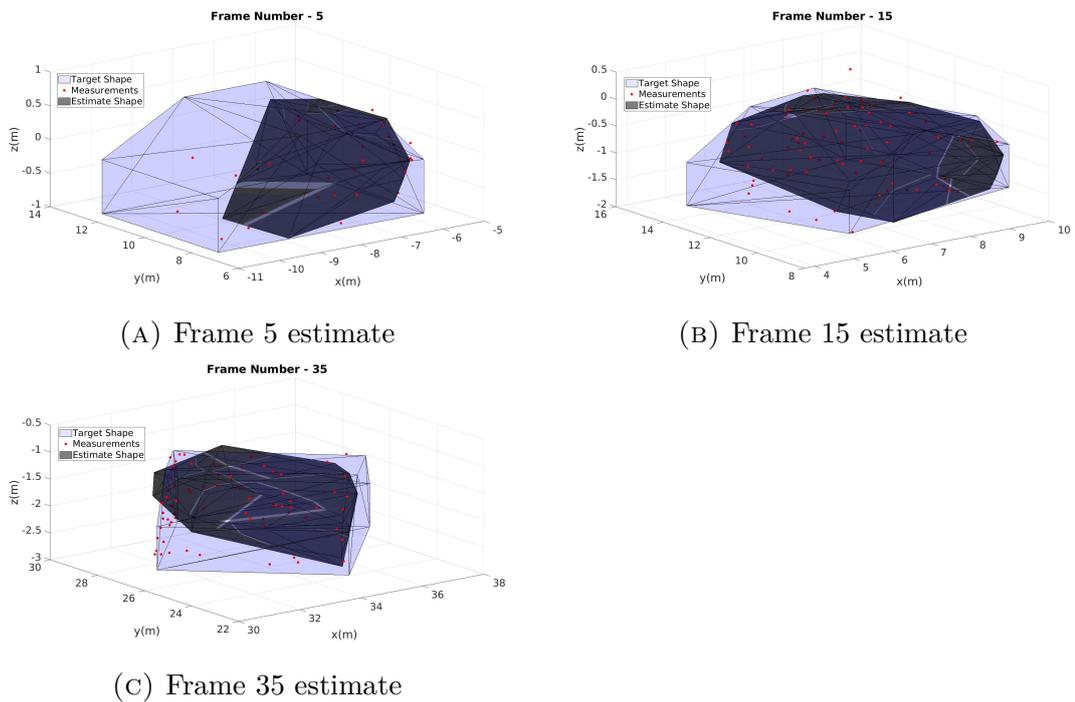


FIGURE 4.11: Estimates of the target shape at different times  $\sigma_x = \sigma_y = \sigma_z = 0.1m$ ,  $\rho = 2m^{-2}$  and  $\rho_C = 0.1m^{-3}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using proposed algorithm is in black.

The results show that proposed algorithm is able to track the target even with low average number of measurements from the target. The performance of the algorithm decreases with decrease in the average number of measurements from the target. Figure 4.11, frame 35 shows that the sample estimate shape does not match the target shape entirely. If the density of measurements is comparable to clutter, new faces might not be added since the clusters are rejected as clutter. The failure to add new faces further affects the algorithm since the measurements that are supposed to be associated to new faces are now used in the estimation of existing faces leading to wrong data association.

Similar to the previous scenario, the Gaussian Process is unable to handle the self-occlusion problem.

#### **4.6.5 Different levels of clutter**

The clutter density  $\rho_C$  is varied as  $\{0.05m^{-3}, 0.1m^{-3}, 0.15m^{-3}\}$ , while the other parameters are kept fixed,  $\sigma_x = \sigma_y = \sigma_z = 0.1m$ , and  $\rho = 0.1m^{-2}$ . Figures 4.12,4.13,4.14 show the RMSE of the center, RMSE of the kinematics and the IoU results. The snapshots of the ground truth, measurements and the estimates at different frames for a sample run is shown in Figure 4.15.

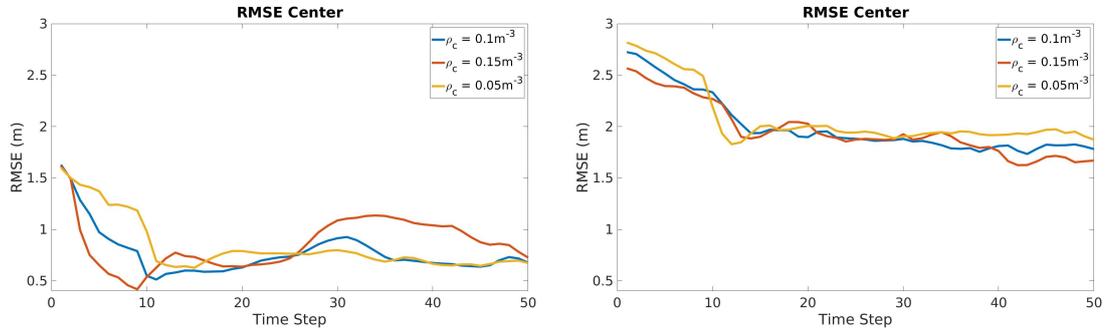


FIGURE 4.12: RMSE center for different levels of clutter. Results for the proposed approach are on the left and the GP results are on the right

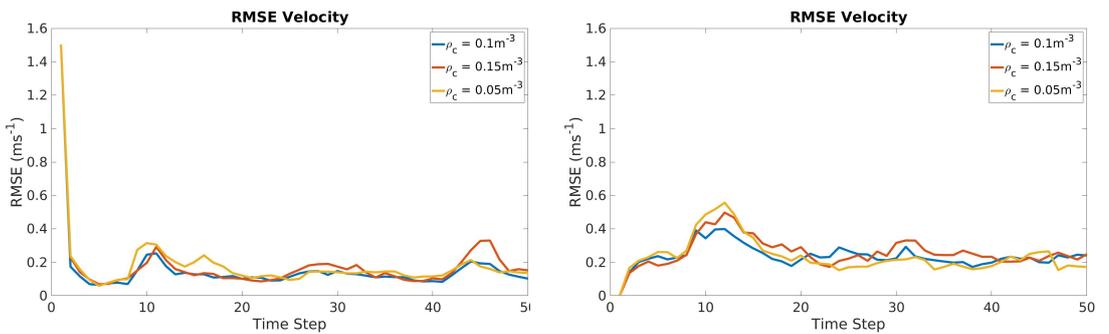


FIGURE 4.13: RMSE velocity for different levels of clutter. Results for the proposed approach are on the left and the GP results are on the right

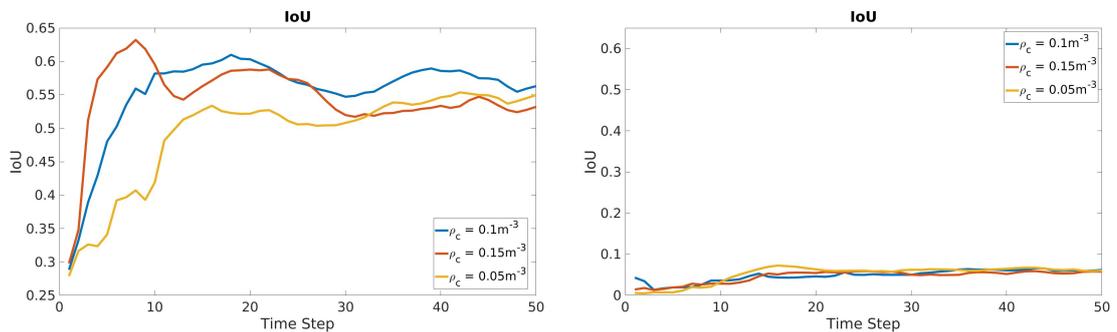
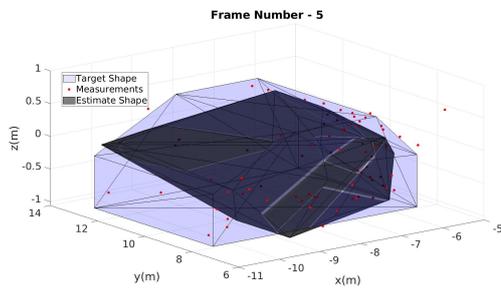
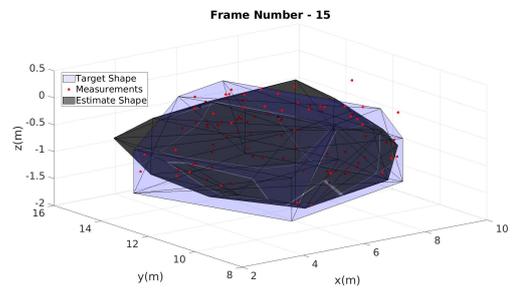


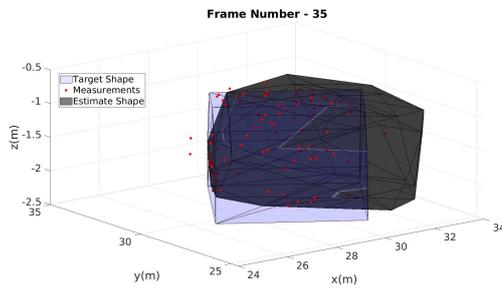
FIGURE 4.14: IoU results for different levels of clutter. Results for the proposed approach are on the left and the GP results are on the right



(A) Frame 5 estimate



(B) Frame 15 estimate



(C) Frame 35 estimate

FIGURE 4.15: Estimates of the target shape at different times  $\sigma_x = \sigma_y = \sigma_z = 0.1m$ ,  $\rho = 3m^{-2}$  and  $\rho_C = 0.15m^{-3}$ . The true target shape is given in blue, the measurements at the current time are given in red. The estimate using proposed algorithm is in black.

The proposed algorithm is able to handle varying clutter densities while estimating the target shape. The snapshot for frame 5 in Figure 4.15 shows that the algorithm estimating a part of the target that is not visible. This is due to higher clutter density and leads to apparent improvement in performance. However this is a drawback of the algorithm and only the observable parts of the target should be estimated. Frame 35 shows the biased nature of the estimate for the sample run in a clear manner. While the estimate shape matches the target shape where the measurements are generated, it fails to maintain the shape that are currently not visible but were visible in previous frames.

## **4.7 Conclusions and future work**

The problem of tracking a maneuvering extended target in 3D with convex polytope shape was discussed. Clutter as well as measurements from the interior of the target were considered. The target maneuvers were modeled by modifying the CT (coordinated Turn) dynamics for extended targets. ET-PMHT equations were derived using linearization to handle the nonlinear dynamics. The effect of clutter was included in the ET-PMHT equations and face management is used to handle clutter as well. Self-occlusion was considered and face initialization was used to add new faces for parts of the target that are visible for the first time. Simulations were performed for different scenarios, and the performance of the proposed algorithm was compared with the 3D Gaussian Process.

The convex polytope constraint on the target shape is a limitation of the current work. While the ET-PMHT equations support arbitrary shapes with appropriate

triangulation, a holistic algorithm to track non-convex shapes should be considered for future work. The effect of different clutter distributions needs to be investigated as well. The extension of the algorithm to a multiple extended target scenario, while considering mutual occlusion, can be attempted as part of future work.

## 4.8 Appendix - Derivation of the Jacobian of the transition function

The Jacobian of the transition function can be derived as

$$\left. \frac{\partial f_t(x(t))}{\partial x(t)} \right|_{x(t)=\hat{x}(t)} = \begin{bmatrix} F_R^{N_X} (I - V^{N_X} M^{N_X}) + V^{N_X} M^{N_X} & V^{N_X} F_{pv} & \nabla_{\Omega_{CT}} X_S(t+1) & \nabla_{\Omega_R} X_S(t+1) \\ 0_{5 \times 3N_X} & F_{vv} & \nabla_{\Omega_{CT}} v(t+1) & \nabla_{\Omega_{CT}} v(t+1) \end{bmatrix} \quad (4.68)$$

where

$$\nabla_{\Omega_{CT}} X_S(t+1) = V^{N_X} \begin{bmatrix} v_x \left[ \frac{T \cos(\Omega_{CT} T)}{\Omega_{CT}} - \frac{\sin(\Omega_{CT} T)}{\Omega_{CT}^2} \right] + v_y \left[ -\frac{T \sin(\Omega_{CT} T)}{\Omega_{CT}} + \frac{1 - \cos(\Omega_{CT} T)}{\Omega_{CT}^2} \right] \\ v_x \left[ \frac{T \sin(\Omega_{CT} T)}{\Omega_{CT}} - \frac{1 - \cos(\Omega_{CT} T)}{\Omega_{CT}^2} \right] + v_y \left[ \frac{T \cos(\Omega_{CT} T)}{\Omega_{CT}} - \frac{\sin(\Omega_{CT} T)}{\Omega_{CT}^2} \right] \\ 0 \end{bmatrix} \quad (4.69)$$

$$\nabla_{\Omega_R} X_S(t+1) = \nabla_{\Omega_R} F_R^{N_x} (I - V^{N_x} M^{N_x}) X_S(t), \quad \nabla_{\Omega_R} v(t+1) = \begin{bmatrix} 0_{4 \times 1} \\ 1 \end{bmatrix} \quad (4.70)$$

$$\nabla_{\Omega_{CT}} v(t+1) = \begin{bmatrix} v_x(-\sin(\Omega_{CT}T)) + v_y(\cos(\Omega_{CT}T)) \\ v_x(-\cos(\Omega_{CT}T)) + v_y(-\sin(\Omega_{CT}T)) \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad (4.71)$$

In the limiting case, as  $\Omega_{CT} \rightarrow 0$ ,

$$\nabla_{\Omega_{CT}} X_S(t+1) = \begin{bmatrix} -\frac{1}{2}T^2 v_y & \frac{1}{2}T^2 v_x & 0 \end{bmatrix}^T \quad (4.72)$$

## Bibliography

- [1] Johann Wolfgang Koch. Bayesian approach to extended object and cluster tracking using random matrices. *IEEE Transactions on Aerospace and Electronic Systems*, 44(3):1042–1059, 2008. doi: 10.1109/TAES.2008.4655362.
- [2] Marcus Baum and Uwe Hanebeck. Tracking an extended object modeled as an axis-aligned rectangle. pages 2422–2427, 01 2009.
- [3] Shishan Yang and Marcus Baum. Tracking the orientation and axes lengths of an elliptical extended object. *IEEE Transactions on Signal Processing*, 67(18):4720–4729, 2019. doi: 10.1109/TSP.2019.2929462.

## BIBLIOGRAPHY

---

- [4] Jian Lan and X. Rong Li. Tracking of maneuvering non-ellipsoidal extended object or target group using random matrix. *IEEE Transactions on Signal Processing*, 62(9):2450–2463, 2014. doi: 10.1109/TSP.2014.2309561.
- [5] Ben Liu, Ratnasingham Tharmarasa, Rahim Jassemi, Daly Brown, and Thia Kirubarajan. Extended target tracking with multipath detections, terrain-constrained motion model and clutter. *IEEE Transactions on Intelligent Transportation Systems*, 22(11):7056–7072, 2021. doi: 10.1109/TITS.2020.3001174.
- [6] Niklas Wahlstrom and Emre Ozkan. Extended target tracking using gaussian processes. *IEEE Transactions on Signal Processing*, 63(16):4165–4178, 2015. doi: 10.1109/TSP.2015.2424194.
- [7] Murat Kumru and Emre Ozkan. Three-dimensional extended object tracking and shape learning using gaussian processes. *IEEE Transactions on Aerospace and Electronic Systems*, 57(5):2795–2814, oct 2021. doi: 10.1109/taes.2021.3067668. URL <https://doi.org/10.1109/2Ftaes.2021.3067668>.
- [8] Xu Tang, Mingyan Li, Ratnasingham Tharmarasa, and Thiagalingam Kirubarajan. Seamless tracking of apparent point and extended targets using gaussian process pmht. *IEEE Transactions on Signal Processing*, 67(18):4825–4838, 2019. doi: 10.1109/TSP.2019.2932873.
- [9] Yunfei Guo, Yong Li, Ratnasingham Tharmarasa, Thiagalingam Kirubarajan, Murat Efe, and Bahadir Sarikaya. Gp - pda filter for extended target tracking with measurement origin uncertainty. *IEEE Transactions on Aerospace and Electronic Systems*, 55(4):1725–1742, 2019. doi: 10.1109/TAES.2018.2875555.

## BIBLIOGRAPHY

---

- [10] Antonio Zea, Florian Faion, Marcus Baum, and Uwe D. Hanebeck. Level-set random hypersurface models for tracking nonconvex extended objects. *IEEE Transactions on Aerospace and Electronic Systems*, 52(6):2990–3007, 2016. doi: 10.1109/TAES.2016.130704.
- [11] Hauke Kaulbersch, Jens Honer, and Marcus Baum. A cartesian b-spline vehicle model for extended object tracking. In *2018 21st International Conference on Information Fusion (FUSION)*, pages 1–5, 2018. doi: 10.23919/ICIF.2018.8455717.
- [12] Karl Granstrom, Marcus Baum, and Stephan Reuter. Extended object tracking: Introduction, overview and applications, 2017.
- [13] Prabhanjan Mannari, Ratnasingham Tharmarasa, and Thiagalingam Kirubarajan. Extended target tracking under multitarget tracking framework for convex polytope shapes. *preprint on SSRN*, 2022. doi: <http://dx.doi.org/10.2139/ssrn.4290822>.
- [14] Prabhanjan Mannari, Ratnasingham Tharmarasa, and Thiagalingam Kirubarajan. 3d extended target tracking using et-pmht for convex polytope shapes with partial visibility. *Ready to be submitted to a journal*, 2023.
- [15] Y. Bar-Shalom and X.R. Li. *Multitarget-multisensor Tracking: Principles and Techniques*. Yaakov Bar-Shalom, 1995. ISBN 9780964831209.
- [16] Y. Bar-Shalom, P.K. Willett, and X. Tian. *Tracking and Data Fusion: A Handbook of Algorithms*. YBS Publishing, 2011. ISBN 9780964831278. URL <https://books.google.ca/books?id=2a0iuAAACAAJ>.

## BIBLIOGRAPHY

---

- [17] Han X. Gaetjens Samuel J. Davey. *Track-Before-Detect Using Expectation Maximisation*. Springer Singapore. ISBN 978-981-13-3971-4. URL <https://link.springer.com/book/10.1007/978-981-10-7593-3>.
- [18] Yaakov Bar-Shalom, X.-Rong Li, and Thia Kirubarajan. *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*. 01 2004. ISBN 047141655X. doi: 10.1002/0471221279.ch11.

# Chapter 5

## Conclusions and Future Work

### 5.1 Conclusions

The thesis discusses the problem of extended target tracking, particularly for convex polytope shapes. The algorithms proposed are able to effectively track 2D and 3D convex polytope shaped targets, considering self-occlusion as well.

An extended target tracking framework was proposed in Chapter 2 and a 2D convex hull shape descriptor was used to develop an algorithm to track 2D convex polytope shapes. Nearest neighbour association was used to determine the source location for the measurements and Kalman filtering was used for joint shape and kinematics estimation. The correctness of the framework was verified using simulations and the performance was compared with the 2D Gaussian Process for different scenarios.

ET-PMHT was derived in Chapter 3 using the framework proposed in Chapter 2 to extend the problem to 3D. The 3D convex polytope shapes were described using

a Delaunay triangulation. The technique of using triangular meshes to describe the shapes of objects is widely used in engineering, especially in CAD (Computer Aided Design) applications. The ET-PMHT equations support arbitrary shapes with appropriate triangulation, however the overall algorithm is restricted to convex polytope shapes.

ET-PMHT is further extended in Chapter 4 to include target maneuvers which occur frequently in practice. A co-ordinated turn model is used to model the maneuvers and is handled in the filtering step using linearization. Clutter as well as measurements from the interior of the target are supported in this work.

The proposed algorithms for 3D extended targets are compared with the 3D Gaussian Process for appropriate scenarios and the performance was quantified using RMSE of the center, RMSE of the velocity and Intersection over Union (IoU) metrics.

### **5.1.1 Challenges and Approaches**

A few of the major challenges faced in the research and the approaches to address them are listed below.

- Development of the Extended target tracking framework - Consider a highly simplified scenario where the kinematics of the target are known perfectly and there is no measurement noise, but the source locations of the measurements from the target shape are unknown. For a convex polytope shape in such a scenario, the total shape can be estimated by successively taking the

convex hull at the current time step, using the predicted shape from the previous time step using the known kinematics, and the current measurements. This technique works even in presence of self-occlusion.

In a more realistic scenario with joint shape and kinematics uncertainty as well as measurement noise, a measurement equation needs to be developed that includes self-occlusion. The measurements generated from a region of the target should be used to estimate only that local region of the target and not affect the entire shape. When faces of the target that were not visible previously become visible over time, the measurements from these faces provide 'new information' about the shape. In light of these issues, the extended target tracking framework was developed. The point multitarget framework uses the association step to group the measurements for each target (sometimes even probabilistically). This inspires the division of the target shapes into multiple faces to be treated as distinct targets and then the measurements can be associated to only the local regions. Track initialization is used in the MTT framework to handle measurements from 'new targets'. Similarly in the proposed framework, new faces can be added to the target shape using measurements from parts of the target that were not visible previously.

- Delaunay Triangulation - The initial idea for the 3D convex polytope shape (Chapter 3 and Chapter 4) was using different polygonal faces (not necessarily just triangles). However the existing convex hull algorithms return a Delaunay triangulation instead. This proved to be more natural since a unique measurement equation can be developed for each face in terms of

its vertices which is not possible for faces with more edges. Further, this method of representing shapes using a triangular mesh is used in various fields of engineering, particularly for simulations using the Finite Element Method.

- Handling the nonlinearity of maneuvers - It can be seen in Chapter 4 that the target state has only one parameter to indicate the turn rate as well as the rotation, but the estimate state has two parameters, one to indicate the turn rate and one to indicate the rotation. Simulations were performed with a single turn rate/rotation parameter in the estimate state and the estimate failed to capture the rotation completely. One of the reasons may be that the turn rate maybe unobservable due to high frame rate and the size of the target, while the rotation is clearly observed. It needs to be investigated further to point out the exact issue in this formulation. The state was then extended to include a separate parameter for rotation and turn rate, which yielded much better performance.

In Chapter 4, Section 4.4.3, the total linear system for one step update was derived. The simulations with the total linear system failed and gave erroneous results. The exact issue to fix the linear system is unknown and can be investigated further. Possibly methods such as preconditioning can be used if the system is ill-conditioned. However in the current work, the problem was split into two sub-problems of updating the kinematics while keeping the shape fixed in Section 4.4.4, and then estimating the shape in Section 4.4.5.

## **5.2 Future Work**

The current work in the thesis is focused on convex polytope shapes and extension to non-convex shapes should be considered in the future. The ET-PMHT equations developed in the thesis support arbitrary shape triangulations including shapes with holes and disjoint faces for joint association and filtering. A coherent algorithm to track non-convex shapes in the presence of self-occlusion and clutter needs to be investigated.

The framework proposed in the thesis supports various multitarget techniques for each block, to be modified for extended targets accordingly. Different shape models, such as the Gaussian Process, and association/filtering techniques, such as PDA and Probabilistic Hypothesis Density (PHD) filter can be used under the framework and their performance can be compared with the proposed algorithm.

Uniform distribution of clutter considered in the thesis may not reflect the realistic scenario in all cases. For example, ground clutter occurs only from below the target and is therefore biased, or clutter from background objects will be correlated across frames. Handling different clutter distributions and analysis on real data can be considered as part of future work.

Currently, only a single extended target is considered with self-occlusion. The single extended target tracking framework proposed can be improved to handle multiple extended targets with mutual occlusion in a hierarchical manner. In reality, the motion of an extended target is influenced by other targets, for example, vehicular traffic on a road. Such effects also need to be considered in a multiple

extended target tracking scenario. Another peculiar effect is that of merge and split, i.e. multiple extended targets merged into a single one or a single target splitting into multiple targets. Such a scenario is common in practice while tracking passengers entering or exiting a vehicle. Including these effects into the multiple extended target framework will lead to more robust and realistic algorithms.