MULTI-TARGET TRACKING WITH UNCERTAINTY IN THE PROBABILITY OF DETECTION

A Thesis

Submitted to the Faculty

of

Purdue University

by

Rohith Reddy Sanaga

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Aeronautics and Astronautics

August 2019

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF THESIS APPROVAL

Dr. Carolin Frueh, Chair

School of Aeronautics and Astronautics

Dr. Kathleen Howell

School of Aeronautics and Astronautics

Dr. Inseok Hwang

School of Aeronautics and Astronautics

Approved by:

Dr. Weinong Wayne Chen

Head of the Graduate Program

Dedicated to my Mom, Sridevi Chegireddy

and

Dad, Srinivasulu Sanaga

ACKNOWLEDGMENTS

Firstly, I would like to thank my mom and dad for giving me the opportunity to follow my dreams and also instilling the importance of education in me. I wouldn't be where I am today if not for your sacrifices. You have continuously supported me and provided me confidence in the most difficult of times and I am eternally grateful for that.

Next, I would like to extend my immense gratitude to my advisor, Dr. Carolin Frueh, for believing in me and agreeing to guide me through this research. Your suggestions and feedback were very valuable to me in completing this research. I would also like to thank my committee members, Dr.Kathleen Howell and Dr.Inseok Huang, for their feedback and support.

I am also grateful for the support from my friends at Purdue. The feedback that I obtained from the members of Space Information Dynamics group was very helpful to me. I would like to thank Sanjit, Sabareesh, Surabhi, Clément for continuously supporting me through this period. Lastly, I would like thank the Department of Computer Science at Purdue for providing me with a Graduate Assistantship to overcome my financial obligations at Purdue.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Page

Figure Page

ABBREVIATIONS

SSA	Space Situational Awareness
MTT	Multi-Target Tracking
SHC	Single-Hypothesis Correlation
MHC	Multi-hypothesis Correlation
RFS	Random Finite Set(s)
IOD	Initial Orbit Determination
p.g.fl(s)	probability generating functional(s)
PHD	Probability Hypothesis Density
CPHD	Cardinalized Probability Hypothesis Density
GM	Gaussian Mixture
EK	Extended Kalman
p_D	Probability of Detection
OSPA	Optimal Subpattern Assignment
w.r.t	with respect to

ABSTRACT

M.S.A.A, Purdue University, August 2019. Multi-Target Tracking with uncertainty in the probability of detection. Major Professor: Professor Carolin Frueh.

The space around the Earth is becoming increasingly populated with a growth in number of launches and proliferation of debris. Currently, there are around 44,000 objects (with a minimum size of 10cm) orbiting the Earth as per the data made publicly available by the US strategy command (USSTRATCOM). These objects include active satellites and debris. The number of these objects are expected to increase rapidly in future from launches by companies in the private sector. For example, SpaceX is expected to deploy around 12000 new satellites in the LEO region to develop a space-based internet communication system. Hence in order to protect active space assets, tracking of all the objects is necessary. Probabilistic tracking methods have become increasingly popular for solving the multi-target tracking problem in Space Situational Awareness (SSA). This thesis studies one such technique known as the GM-PHD filter, which is an algorithm which estimates the number of objects and its states when non-perfect measurements (noisy measurements, false alarms) are available. For Earth orbiting objects, especially those in Geostationary orbits, ground based optical sensors are a cost-efficient way to gain information.In this case, the likelihood of gaining target-generated measurements depend on the probability of detection (p_D) of the target.An accurate modeling of this quantity is essential for an efficient performance of the filter. p_D significantly depends on the amount of light reflected by the target towards the observer. The reflected light depends on the relative position of the target with respect to the Sun and the observer, the shape, size and reflectivity of the object and the relative orientation of the object towards Sun and the observer. The estimation of the area and reflective properties of the object is in

general, a difficult process. Uncontrolled objects, for example, start tumbling and no information regarding the attitude motion can be obtained. In addition, the shape can change because of disintegration and erosion of the materials. For the case of controlled objects, given that the object is stable, some information on the attitude can be obtained. But materials age in space which changes the reflective properties of the materials. Also, exact shape models for these objects are rare. Moreover,, area can never be estimated with optical measurements or any other measurements, as it is always albedo-area i.e., reflectivity times area that can be measured.

The purpose of this work is to design a variation of the GM-PHD filter which accounts for the uncertainty in p_D as the original GM-PHD filter designed by Vo and Ma ([1]) assumes p_D as a constant. It is validated that the proposed method improves the filter performance when there is an uncertainty in area(hence uncertainty in p_D) of the targets. In the tested cases, the uncertainty in p_D was modeled as an uncertainty in area while assuming that the targets are spherical and that the reflectivity of the targets is constant. It is seen that a model mismatch in p_D affects the filter performance significantly and the proposed method improves the performance of the filter in all cases.

1. INTRODUCTION

The space around Earth is becoming more crowded with increasing launches and the proliferation of debris. Currently, there are around 44000 Earth-orbiting objects of a minimum size of 10 centimeters (refer [2]). These objects are maintained in a publicly available catalog maintained by the US Strategic Command (USSTRATCOM). These objects consist of the active satellites and also space debris. Debris include the inactive/dead satellites, mission related objects like the rocket upper stages and also the objects created from the fragmentation of the materials of satellites and the objects created from collisions. By definition, space debris consists of only human-made objects. The number of objects is expected to increase rapidly in the coming years. For example, SpaceX is expected to deploy around 12000 objects in the Low Earth Orbits by mid 2020's (refer [3]). Hence, sufficient information on the state (position and velocity) of these objects is needed to protect and manage active space assets. This can be achieved by *tracking* which is the process of estimating a required quantity (like state, number of objects in a region) from measurements (like radar measurements,optical measurements). Collisions not only result in losses of the space assets but also increase the debris in space because of fragmentation of the objects involved in the collision. Hence, tracking is an significant part of Space Situational Awareness (SSA).

There are two major challenges in tracking in SSA. The first one is the problem of detection of new objects and the second one is updating the orbits of known, catalogued objects. Hence, a tracking algorithm uses the measurements to estimate any existing target's state in a surveillance state and also define any new object entering the scene which was previously undetected. The term cataloging is sometimes used in the SSA community and is synonymous with tracking in this context. Tracking involves various sub-parts which include measurement collection, initial orbit deter-

mination (IOD), data association, clutter rejection and track identification. The first step in tracking is to process the images obtained to extract the measurements. IOD is the process of getting a initial estimate of a target's orbit. This is used in tracking algorithms to estimate an initial state of a target. Clutter rejection is done to discern the target-generated measurements from the false measurements (Example:Measurements generated because of cosmic rays). Data association is the process of associating a measurement to an existing track/object or a new object which helps to improve/update the state estimate of the target. Track labelling is done for the unique identification of a target. An efficient tracking algorithm should provide all of the above mentioned components. In single target tracking, only one object needs to be tracked. This makes the data association step easier since only the clutter has to be identified instead of clutter and the specific target-generated measurement. Multi-Target Tracking methods, in comparison, deal with the tracking of multiple objects simultaneously or quasi-simultaneously. Also, MTT algorithms need to model any new objects entering a scene which is not the case with single target tracking. As discussed before, in SSA, tracking existing objects and new, previously undetected objects is necessary. Hence, the following section describes MTT and the various approaches pursued to develop MTT techniques.

1.1 Multi-target Tracking (MTT)

MTT is done to obtain or estimate the number of objects and their states in a surveillance scene from a set of observations which include noise and false alarms/clutter measurements. Data fusion is the process of integrating multiple data sources to produce consistent, accurate and useful information about the entity which is being observed. The MTT and data fusion literature has many techniques. [4] provides a taxonomy of such methods and also details and compares the performance of various MTT algorithms. Also, articles such as [5], [6], [7] and [8] and reference texts such as [9], [10] and [11] provide further insight into MTT algorithms.The MTT problem

has been explored by two major approaches, namely, track-based approaches and population based approaches. The following sections give a brief overview of these approaches.

1.1.1 Track-Based Approaches

As the name suggests these approaches keep track of individual targets separately. Hence, this approach includes the data-association problem i.e., the measurements are explicitly associated with the targets. Since the number of measurements and the targets are in general not same (because of false alarms, missed detections), all the possible measurement-to-track combinations are considered. Hence, an important component considered here is the *Global Association Likelihood* (GAL) which is a numerical quantity which quantifies how likely a combination occurs. The simplest type of filter is the *Single-Hypothesis Correlation* (SHC) filter (for detailed discussion, refer [12]). Based on the GAL, a SHC filter selects one measurement-to-track association at every time step and estimates tracks for every target in the scene using the measurements. The performance of SHC filters declines when the targets are closely spaced. This is because SHC filter decides a particular measurement-to-track association as a correct one which might not be true when the targets are closer. *Multiple-Hypothesis Correlation* (MHC) mitigate this by keeping track of all the measurement-to-track associations i.e., in addition to keeping information on the tracks, it also keeps track of different hypotheses. More precisely, at any time step, every possible measurement-to-track association is stored and corresponding probability is assigned to it depending on GAL. These probabilities are then used in estimating the number of targets and their states. MHC filters answer the constraints of SHC filters but are computationally more complex and expensive. The Composite-hypothesis Correlation (CHC) filters reduces the computation load by merging SHC and MHC. In essence, CHC filters select a composite hypothesis at every time step.A measurement is assumed to be associated to some extent to every target (detailed discussion

in Chapter 10, [12]) i.e., a probability density is assigned to this using the hypothesis density and then the data update step is carried out by taking a weighted average of all the measurement-to-track associations . The most well known filter of this type is the *Joint Probabilistic Data Assosciation* (JPDA) filter, introduced in [13], which is used when the number of objects are known. Track-based approaches usually rely on a heuristic approach (track pruning, creating, merging) and must be tuned according to the filtering problem.

1.1.2 Population-Based Approaches

The second approach known as the Population-based approach considers the group of targets as a single entity and hence doesn't explicitly solve the data association problem. As a result, a specific target track cannot be naturally estimated.Even though some information is lost regarding the individual targets this approach is comparatively less complex. In this approach, the information of the target states is stored in a single random object known as Random Finite Sets (RFS) whose size (number of targets) and also the elements (target states) are random. Also, in this formulation the measurements consisting of the target-generated measurements and false measurements (clutter) are also modeled as RFS. Modeling the measurements and the states of the targets as sets allows the Multi-target filtering problem to be looked in a Bayesian filtering framework. This can be done by using the concepts and formulations in FISST (Finite Sets Statistics). The Multi-target Bayes filter can then be obtained as a generalisation of the single target Bayes filter. This thesis gives a systematic derivation of Multi-Target Bayes filter in the following chapters. This approach was formally developed by Mahler in [12]. However, in many cases, the Multi-target Bayes filter is computationally intractable (detailed discussion in chapter 3). To alleviate this limitation, the PHD (probability density hypothesis) filter (discussed in chapter 4) was developed which is a less computationally expensive version of the Multi-target Bayes filter. The PHD filter propagates only the first-

order statistical moment of the RFS which is a quantity analogous to the mean of a random variable. Even though the complexity is increased the filter involves multiple integrations which in general have no closed form solutions. To gain a closed form solution to the filter additional assumptions are made and the GM-PHD filter was designed by [1]. These assumption can be thought of an extension of the Kalman filter to the multi-target regime. This thesis focuses on the application of GM-PHD filter in SSA.

1.2 MTT Applications in SSA

Both the track-based and the population-based approaches have been applied in the field of SSA. Frueh et al. [14] and Frueh [15] propose cosmic filters, which link multiple non-resolved images belonging to a same object in the presence of cosmic rays. The linked images are then used for initial orbit determination (IOD). Cosmic rays provide false detections and are indistinguishable from the object detections. To filter them out, multiple images are used to isolate the object images since the cosmic rays are singular events and non-regular (refer [15]). The cosmic filters have a kind of track based approach since there is explicit data association. Kelecy et al. [16], propose a MHC filter to track High Area-to-Mass Ratio (HAMR) objects in the vicinity of Geostationary Orbits. This approach is used since the optical surveys from this region result in large number of uncorrelated tracks and it is showed that the MHC filter is an efficient approach to track these objects. Singh et al. [17] propose a MHT filter to resolve uncorrelated tracks and uncorrelated optical observations.
Population-based approach or the FISST approach has also been used in the SSA area. DeMars et al. [18], Gehly et al. [19], Jones et al. [20], [21], Cheng et al. [22], [23] use random finite sets approach to find solutions in the SSA scenario. DeMars et al. [18], propose a filter similar to the GM-PHD filter but instead of having Gaussian assumptions in the PHD filter, the Gaussian assumptions are introduced directly in the Multi-Bayes filter to reduce information loss resulting from the first order approx-

imation in the PHD filter. Gehly et al. [19] utilizes the CPHD (Cardinalized PHD) filter to propose a method to track GEO objects when there is a Field of View constraint for the measurements. CPHD filter is an extension of PHD filter in the sense that it propagates the intensity function (PHD) and also the cardinality distribution of the targets. Jones et al. [20] propose a method to initiate new objects entering a surveillance scene in the CPHD filter. This method uses the measurements to initiate and track unknown targets while simultaneously tracking the known objects. Jones et al. [21] propose the multi-Bernoulli filter which allows for the identification of the individual targets. This method alleviates the limitation of the population based approaches which are not designed to track individual targets. Cheng et al. [22] compares the performance of PHD and CPHD filters in the scenario of space object tracking. The CPHD filter is shown to be more accurate in estimating the intensity function as well as the number of objects but with more computational cost. Cheng et al. [23] investigates the viability of the application of GM-PHD filter to space object tracking. It proposes a filter which accounts for the non-linear effects obtained because of long term propagation of the orbits. As mentioned previously, this thesis focuses on the GM-PHD filter. The following section details the assumptions made to obtain the GM-PHD filter.

1.3 Assumptions of GM-PHD Filter

The following points describe the major assumptions of the GM-PHD filter (refer [1]).

- **A1**:*Each target moves and generates measurements independently of each other.* In SSA, Target independence is a valid assumption for most of the cases. This assumption is violated only in the case of collision event and in the case of decaying objects and drag wake effects ([24]). Measurement independence is also a good assumption in the case of optical measurements (refer [24] for detailed explanation). This thesis will work with optical measurements.

- **A2**:*Clutter is Poisson and independent of target-generated measurements.*
  A detailed explanation of this assumption is given in chapter 4. However, this assumption is valid in SSA. The independence is established by the fact that the physical processes leading to the target-generated and clutter measurements are independent. Clutter is usually a result of cosmic ray events which are a very good example of a Poisson process.

- **A3**:*Predicted multiple-target RFS is Poisson.*
  This assumption too will be discussed in detail in chapter 4. Assuming the predicted multi-target RFS to be Poisson implies that the actual number of targets can be assumed to "jump". This is only the case when collisions are involved and in general, the actual population, depending on a surveillance scene, is not highly variable. Hence, this assumption cannot be readily made in SSA.

- **A4**:*Probability of survival (p_S) and Probability of detection (p_D) are constant.*
  Constant p_S is a good assumption for the case of SSA (refer [24]). However, constant p_D assumption cannot be readily made in SSA regime. This thesis focuses on finding an alternative version of the GM-PHD filter by not making this assumption.

The following section details why constant p_D is not a valid assumption in SSA scenario and lists out the specific objectives of this research.

1.4 Objective

As will be seen in chapter 4 and as outlined above, GM-PHD filter in its basic form assumes that p_D is a constant and is not dependent on the state of the target. This assumption was used to obtain a closed form solution for the PHD filter and is assumed to be valid in many tracking algorithms [25], [10], [26], [27], [28]. For example, in [27] the GM-PHD filter is used for visual people tracking where people

are tracked from a set of images obtained from a video . The probability of detection here can be assumed to be a constant depending on the quality of the camera used to capture the video since whoever is in the Field of View (FOV) of the camera will surely show up in the image. However, this assumption is not a very good one in the space environment.

For Earth orbiting objects, especially those in Geostationary orbits, ground based optical sensors are a cost-efficient way to gain information ([29]). CCD (Charged couple device) sensors are usually used to take the optical observations. In order to generate a measurement, the target must reflect sufficient light to the sensor. The reflected light received by the sensor depends on the relative position of the object with the observer and the illumination source (Sun), the reflective properties of the object and the attitude motion of the object. Light is also attenuated through transition through the atmosphere in ground-based observations. This is known as atmospheric extinction. The sensor properties also determine the probability of detection. For example, the aperture size decides how much light is taken in by the sensor. Increasing it would increase p_D, but it could also increase the clutter measurements. Hence, modeling p_D as a constant although it is highly variable, creates a model mismatch. [24] shows that a model mismatch in probability of detection affects the performance of a simplified cardinality-only PHD filter. The present work shows the effects of model mismatch in p_D in the full PHD filter.

As mentioned above, p_D depends on the reflective area of the target. Moreover, over 93 percent of the catalogued objects are uncontrolled and non-operational space objects, so-called space debris. Since these objects are uncontrolled, the attitude motion of these objects cannot be estimated accurately and hence their reflective area cannot be modeled accurately. Also, over time, there is degradation of the surface materials which changes the reflective properties of the object. The estimation of the shape of these objects in itself is a very complex process. In the case of the controlled objects, the reflected light mostly comes from the solar panels. In some cases, for example, in the case of controlled satellites in high orbits, the solar panels always point to the

Sun and the reflected light comes from the body of the satellite. This makes the prediction of the reflective area very difficult. Moreover, p_D is also time-dependent. For example, consider an object which is initially between the Sun and the observer. Here, since there is no light reflected to the sensor, p_D should be zero. As the object moves about its orbit, p_D will slowly increase and can also be equal to one depending on the reflected light to the sensor. Therefore, in the case of SSA, p_D can attain its maximum and minimum value within one time period of its orbit.

To summarize the previous paragraph, it can be seen that, in the space environment, p_D is a state (hence time)-dependent and an uncertain quantity. The time dependence is a result of the attitude motion of the object and also the change in relative position of the object with respect to the position of the Sun and the observer. The p_D uncertainty is introduced because of the degradation of materials in space, the unknown shapes (areas) and attitude motion of debris.

Hence, all these factors introduce a level of uncertainty in p_D. This has to be accounted for in the filter because a model mismatch in p_D will affect the filter performance. p_D uncertainty and aslo the state dependence is not considered in any of the previous work mentioned in section 1.2. This work investigates the effects of model mismatch in p_D in SSA and shows that the effects are not negligible. A variation of GM-PHD filter is developed, which models p_D as a state (hence time)-dependent quantity. And finally, a method which incorporates an uncertainty in p_D resulting from an uncertainty in area, into the GM-PHD filter is proposed and validated . The following section gives an overview of the chapters in this thesis.

1.5 Overview

- **Chapter 2**: This chapter introduces the necessary background required to apply FISST. This chapter reviews the basic models used in the single target filtering and defines the single target Bayes filter. The concept of Random

Finite Sets (RFS) is introduced and the necessary calculus required to derive the multi-target Bayes filter is explored.

- **Chapter 3**: A formal derivation of Multi-Target Bayes filter is presented in this chapter. The formulation of the Multi-target Measurement and State space using RFS is shown. Using these, the Multi-target likelihood function and the Multi-Target Markov density function are then defined and Bayes rule is applied to define the Multi-Target Bayes filter.

- **Chapter 4**: This chapter introduces the first-order moment of a Random Finite Set, namely, the Probability Hypothesis Density (PHD). A formal mathematical definition of PHD is presented which is used to derive the PHD filter from the Multi-Target Bayes Filter. The assumptions required for the derivation of the GM-PHD filter are reviewed and the filter equations of the GM-PHD filter are presented.

- **Chapter 5**: This chapter lists the motion models and measurement models used in the simulations. A two-body motion model is assumed for the targets in the space environment and the measurements are assumed to be optical measurements. The Birth model in the space environment used to model the new objects entering the FOV is reviewed. The concept of Admissible region is reviewed for this purpose.

- **Chapter 6**: This chapter provides a analytical expression for p_D of a space object. It is seen that p_D depends on the amount of light reflected by the object and also the sensor properties. A variation of GM-PHD filter which incorporates a state dependent p_D is introduced. The last section proposes a methodology to incorporate an uncertainty in p_D because of an uncertainty in Area in the GM-PHD filter.

- **Chapter 7**: This chapter includes the results of application of the basic GM-PHD filter in SSA. The results obtained by introducing model mismatch in p_D

are displayed and analyzed. The model proposed to include uncertainty in p_D is validated with suitable simulations.

2. BACKGROUND

As was mentioned in the introduction, Random Finite Sets(RFS) are used in the formulation of the multi-target filtering problems.The purpose of using Finite Set Statistics(FISST) is not because RFS are used to model the problem but rather it is because random set techniques provide a systematic toolbox of techniques to address many challenges of multi-target tracking. Fundamental to the development of this thesis is the concept of the Bayes Filter, a well-developed approach to recursive state estimation. The following section reviews the single-target recursive Bayes Filter([9], [30], [31], [32]) and then the Kalman filter is reviewed which is a closed form solution to the single-target recursive bayes filter. This will act as a stepping stone to the concepts of FISST and multi-target filtering.

2.1 Single-target Recursive Bayes Filter

The first step in any filtering problem is to mathematically model the dynamic motion of the target and model the measurement. A general discrete model for the motion of the target can be given by a *Markov transition*:

$$\mathbf{X}_{k+1} = \Phi_k(\mathbf{x}) + \mathbf{V}_k \tag{2.1}$$

where $\mathbf{X}$ is the state of the target and belongs to the vector space $\mathcal{R}^{n_x}$ (n_x is the dimension of the state vector), $\Phi(x)$ is a vector valued function of x, $\mathbf{V}$ is the system noise added to compensate for any modelling errors and k is any time instant. *Markov Transition* means that the state of the body at any time step depends only on the state at the previous time step. Alternatively, the time evolution of the state vector can be described by a *Markov transition density* $f_{k+1|k}(.|.)$ where:

$$f_{k+1|k}(\mathbf{x}_{k+1}|\mathbf{x}_k) \tag{2.2}$$

is the probability that the target has a state $\mathbf{x}_{k+1}$ at t_{k+1} given that he has a state $\mathbf{x}_k$ at t_k.

Correspondingly, the measurement model is defined as:

$$\mathbf{Z}_{k+1} = \eta_{k+1}(\mathbf{x}) + \mathbf{W}_{k+1} \tag{2.3}$$

where $\mathbf{Z}$ is the measurement and belongs to the vector space $\mathcal{R}^{n_z}$ (n_z is the dimension of the measurement vector), $\eta(\mathbf{x})$ is a vector valued function of x, $\mathbf{W}$ is the measurement noise. Alternatively, the measurement model can be described by a *likelihood function* $f_{k+1}(.|.)$ where

$$f_{k+1}(\mathbf{z}|\mathbf{x}) \tag{2.4}$$

is the probability that at time t_{k+1} state $\mathbf{x}$ generates a measurement $\mathbf{z}$. Consider the term $f_{k+1|k}(\mathbf{x}|\mathbf{x}', \mathbf{Z}^k)$. This is known as the state transition density and it denotes the probability density that the state $\mathbf{x}$ occurs at t_{k+1} given the measurement history till t_k i.e., $\mathbf{Z}^k = (\mathbf{Z}_1, ..., \mathbf{Z}_k)$ and the state at $t_k(\mathbf{x}')$. Then the probability density of the state vector conditioned on the measurement history can be given by $f_{k+1|k}(\mathbf{x}|\mathbf{Z}^k)$, which is:

$$f_{k+1|k}(\mathbf{x}|\mathbf{Z}^k) = \int f_{k+1|k}(\mathbf{x}|\mathbf{x}', \mathbf{Z}^k).f_{k|k}(\mathbf{x}'|\mathbf{Z}^k)d\mathbf{x}' \tag{2.5}$$

Proof can be found in [12](page 61). Notice the term $f_{k|k}(\mathbf{x}'|\mathbf{Z}^k)$ in equation 2.5, it gives the probability density of the state at t_k conditioned on the measurement history till t_k. This is known as the *posterior density* or the *filtering density*. It contains all the information about the state vector up to time k in the sense that it encompasses the modelling information and uses all the measurements available at t_k. The posterior density can then be obtained by the *Bayes' rule*. Proof can be found in [12] (page 62).

$$f_{k+1|k+1}(\mathbf{x}|\mathbf{Z}^{k+1}) = \frac{f_{k+1}(\mathbf{z}_{k+1}|\mathbf{x}, \mathbf{Z}^k).f_{k+1|k}(\mathbf{x}|\mathbf{Z}^k)}{f_{k+1}(\mathbf{z}_{k+1}|\mathbf{Z}^k)} \tag{2.6}$$

equation 2.5 is called as the predictor equation and equation 2.6 is called as the corrector equation and together they form the equations for the Bayes Recursive filter given a initial density $f_{0|0}(\mathbf{x})$. The filter then propagates as follows:

$$f_{0|0}(\mathbf{x}|\mathbf{Z}^{(0)}) \rightarrow f_{1|0}(\mathbf{x}|\mathbf{Z}^{(0)}) \rightarrow f_{1|1}(\mathbf{x}|\mathbf{Z}^{(1)}) \rightarrow$$
$$... \rightarrow f_{k|k}(\mathbf{x}|\mathbf{Z}^{(k)}) \rightarrow f_{k+1|k}(\mathbf{x}|\mathbf{Z}^{(k)}) \rightarrow f_{k+1|k+1}(\mathbf{x}|\mathbf{Z}^{(k+1)}) \rightarrow ...$$

The optimal state at every time step can then be obtained from the posterior distribution. The common estimates are expected a posteriori (EAP) estimator and the maximum a posteriori (MAP) estimator [12] (page 63). EAP is the expected value whereas MAP is the state vector with highest probability.

Due to the multiple integrations in the Bayes recursion equations, a closed form solution is difficult to obtain until certain assumptions are made. Though it is possible to find numerical solutions to the above equations they are computationally expensive especially when the dimension of the state space is high. The following section reviews the Kalman Filter which is a closed form solution to the Bayes Recursive filter.

2.2 Kalman Filter

The Kalman filter (refer [32], [12], [33]) is a closed form solution to Bayes Recursion. The Kalman filter assumes that the dynamic model describing the motion and measurement model are linear with additive Gaussian noise.

$$\mathbf{X}_{k+1} = F_k\mathbf{x} + \mathbf{V}_k \tag{2.7}$$

$$\mathbf{Z}_{k+1} = H_{k+1}\mathbf{x} + \mathbf{W}_{k+1} \tag{2.8}$$

where F_k is the transition matrix and H_{k+1} is the observation matrix. V_k and W_k are independent gaussian random variables with zero mean and covariances Q_k and R_k respectively. It is also assumed that the markov transition density and the measurement likelihood are Gaussian and are given by:

$$f_{k+1|k}(\mathbf{x}_{k+1}|\mathbf{x}_k) = \mathcal{N}(\mathbf{x}_{k+1}; F_k\mathbf{x}_k, Q_k) \tag{2.9}$$

$$f_{k+1}(\mathbf{z}_{k+1}|\mathbf{x}_{k+1}) = \mathcal{N}(\mathbf{z}_{k+1}; H_{k+1}\mathbf{x}_{k+1}, R_{k+1}) \tag{2.10}$$

Substituting them in the Bayes recursion equations(2.5,2.6) we obtain the recursive equations for the Kalman filter i.e., if at t_k the posterior density has a Gaussian distribution of the form:

$$f_{k|k}(\mathbf{x}_k|\mathbf{Z}^k) = \mathcal{N}(\mathbf{x}_k; m_k, P_k) \tag{2.11}$$

where the distribution is given by:

$$\mathcal{N}(\mathbf{x}; m, P) = \frac{1}{\sqrt{(2\pi)^k det(P)}} e^{\frac{-1}{2}(\mathbf{x}-m)^T P(\mathbf{x}-m)} \tag{2.12}$$

where $k = n_x$, $det(P)$ is the determinant of P. Then the predicted density at t_{k+1} is also a Gaussian and has the form:

$$f_{k+1|k}(\mathbf{x}_{k+1}|\mathbf{Z}^k) = \mathcal{N}(\mathbf{x}_{k+1}; m_{k+1|k}, P_{k+1|k}) \tag{2.13}$$

where the mean $m_{k+1|k}$ and the co-variance $P_{k+1|k}$ of the predicted state $\mathbf{x}_{k+1|k}$ is given by:

$$m_{k+1|k} = F_k \mathbf{x}_k \tag{2.14}$$

$$P_{k+1|k} = F_k P_k F_k^T + Q_k \tag{2.15}$$

The corrector equation then results in a Gaussian posterior density which is given by:

$$f_{k+1|k+1}(\mathbf{x}_{k+1}|\mathbf{Z}^{k+1}) = \mathcal{N}(\mathbf{x}_{k+1}; m_{k+1}, P_{k+1}) \tag{2.16}$$

where,

$$m_{k+1} = m_{k+1|k} + K(\mathbf{z}_{k+1} - H_{k+1}\mathbf{x}_{k+1|k}) \tag{2.17}$$

$$P_{k+1} = [I - KH_k]P_{k+1|k} \tag{2.18}$$

$$K = P_{k+1|k} H_{k+1}^T S^{-1} \tag{2.19}$$

$$S = R_{k+1} + H_{k+1} P_{k+1|k} H_{k+1}^T \tag{2.20}$$

The matrix $K, \mathbf{z}_{k+1} - H_{k+1}\mathbf{x}_{k+1|k}, S$ are most commonly known as the *Kalman Gain*, *innovation* and *innovation matrix* respectively. It is important to note that the filter

has to be initialized with a Gaussian distribution i.e., $f_{0|0}$ must be Gaussian since it is one of the assumptions. Then the Kalman recursion ensures that the subsequent posterior densities $(f_{1|1}, ..., f_{k|k}, ...)$ are Gaussian. This is an important result of the kalman filter since the posterior distribution at any time instant can be described by mean and co-variance i.e., only the mean and co-variance need to be propagated or updated at each time instant. Although the Kalman filter gives a closed form solution to the Bayes recursive filter it is only applicable when the motion and the measurement models can be represented by linear models and when the filter is initialized with a Gaussian Mixture(i.e., $f_{0|0} = \mathcal{N}(\mathbf{x}; m_{0|0}, P_{0|0})$).
But in any space problem i.e., in the analysis of Earth-orbiting objects the most simple dynamic model is the two-body problem which is non-linear. In order to incorporate this non-linearity an approximate technique called the Extended Kalman Filter [32] is used. This method is briefly described in the next section.

2.3 Extended Kalman Filter

The Extended Kalman Filter (refer [32]) is a first order approximation of the Kalman Filter i.e., the measurement and motion models are linearized using a Taylor series expansion to approximate the predicted and posterior densities as Gaussian distributions. Linearization is important since a Gaussian distribution doesn't remain a Gaussian under non-linear transformation i.e., if $\mathbf{x}_{k|k}$ is Gaussian then under the transformation $\Phi_k(\mathbf{x})$,$\mathbf{x}_{k+1|k}$ will not be a Gaussian if Φ_k is non-linear. Hence Φ_k is linearized to preserve the Gaussian nature. Consider the motion and measurement models described by the equations 2.1 and 2.3 where Φ_k and η_k are non-linear functions of the state vector. Then if at t_k the posterior density is a Gaussian of the form:

$$f_{k|k}(\mathbf{x}_k|\mathbf{Z}^k) = \mathcal{N}(\mathbf{x}_k; m_k, P_k) \tag{2.21}$$

then the predicted density at t_{k+1} is also a Gaussian and has the form:

$$f_{k+1|k}(\mathbf{x}_{k+1}|\mathbf{Z}^k) = \mathcal{N}(\mathbf{x}_{k+1}; m_{k+1|k}, P_{k+1|k}) \tag{2.22}$$

where the mean $m_{k+1|k}$ and the co-variance $P_{k+1|k}$ of the predicted state $\mathbf{x}_{k+1|k}$ is given by:

$$m_{k+1|k} = \Phi_k(\mathbf{x}_k) \tag{2.23}$$

$$P_{k+1|k} = F_k P_k F_k^T + Q_k \tag{2.24}$$

$$F_k = \left.\frac{\partial \Phi_k(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=m_k} \tag{2.25}$$

The Gaussian posterior density is then given by:

$$f_{k+1|k+1}(\mathbf{x}_{k+1}|\mathbf{Z}^{k+1}) = \mathcal{N}(\mathbf{x}_{k+1}; m_{k+1}, P_{k+1}) \tag{2.26}$$

where,

$$m_{k+1} = m_{k+1|k} + K(\mathbf{z}_{k+1} - \eta_{k+1}(\mathbf{x}_{k+1|k})) \tag{2.27}$$

$$P_{k+1} = [I - KH_k]P_{k+1|k} \tag{2.28}$$

$$K = P_{k+1|k} H_{k+1}^T S^{-1} \tag{2.29}$$

$$S = R_{k+1} + H_{k+1} P_{k+1|k} H_{k+1}^T \tag{2.30}$$

$$H_k = \left.\frac{\partial \eta_{k+1}(\mathbf{x})}{\partial \mathbf{x}}\right|_{\mathbf{x}=m_{k+1|k}} \tag{2.31}$$

These equations will be revisited again in the derivation of the PHD filter in the later chapters.

The tools and techniques used in the single-target filtering have been reviewed until this point. The Bayes multi-target recursive filtering equations have the same form as equations 2.5 and 2.6. But before these equations are introduced, the concept of Random Finite sets have to be introduced first. This is because in the single target case, the definition of the probability density and mathematical operations like the differentiation and integration is straightforward. But in multi-target case where sets are used in the modeling of the problem, the concepts like integrating over sets(set integral), differentiation w.r.t a set (set derivative) have to be defined first i.e., calculus of sets or the multi-object calculus has to introduced. The next section provides the example of an application of a random finite set and then formally define a random finite set, set integral , multi-object density function, etc.

2.4 Random Finite Sets

Consider a case when a single sensor is taking observations($\mathbf{z}$) over a certain region of space. The observation consists of the right-ascension(α) and declination(δ) of the object($\mathbf{z} = (\alpha, \delta)$). At any instant of time the sensor collects multiple observations($\mathbf{z}_1, ..., \mathbf{z}_m$) which include the observations from actual targets(number of observations can be any non-negative integer) and clutter observations which are false measurements generated from the sensor itself or can be from background sources in the field of view (Eg : Stars). Also, even though measurement's components have a fixed order(α, δ) the detections ($\mathbf{z}_1, ..., \mathbf{z}_m$) have no inherent physical order. Then the measurement can be represented by a *finite observation set* whose elements are the detections:

$$Z = \{\mathbf{z}_1, ..., \mathbf{z}_m\} = \{(\alpha_1, \delta_1), ..., (\alpha_m, \delta_m)\} \tag{2.32}$$

The measurement space $\mathcal{M}$ is the hyperspace of all finite subsets of an underlying measurement space $\mathcal{M}_0$(here it's $\mathcal{R}^2$). Not only are the components of a individual measurement($\mathbf{z}_i$) random, the number of detections (m) are also random. Thus a measurement is a randomly varying set (Σ) whose elements belong to $\mathcal{M}_0$ i.e., the various samples (instantiations) of Σ are:

$$Z = \emptyset \text{ (no detections)} \tag{2.33}$$

$$Z = \{\mathbf{z}_1\} \text{ (a single detection } \mathbf{z}_1 \text{ is collected)} \tag{2.34}$$

$$Z = \{\mathbf{z}_1, \mathbf{z}_2\} \text{ (two detections } \mathbf{z}_1 \neq \mathbf{z}_2 \text{ are collected)} \tag{2.35}$$

and a sample with m detections is given by:

$$Z = \{\mathbf{z}_1, ..., \mathbf{z}_m\} \text{ (m detections } \mathbf{z}_1 \neq ... \neq \mathbf{z}_m \text{ are collected)} \tag{2.36}$$

Therefore, Σ is an example of a Random Finite Set (RFS) and it's instantiations are random sets whose elements belong to the sensor measurement space $\mathcal{M}_0$. Now that the concept of RFS is well-defined the following sections introduce the concepts of Multi-object Calculus (set functions, set integrals). Set derivatives and set integrals

are necessary to clearly state the multi-object probability distributions and also derive the multi-object Markov densities(analogous to equation 2.2) and multi-target likelihood function(analogous to equation 2.4) from the corresponding multi-object motion models and multi-target measurement models which will be discussed in the following chapters.

2.4.1 Set Function

A set function is a real-valued function $\phi(S)$ of a closed set variable S, whose elements belong to some vector space $\mathcal{Q}_0$. The following set functions are very important in the multi-object filtering literature.

- *Probability-mass function*:If $\mathbf{Q}$ is a random vector on $\mathcal{Q}_0$, then:

$$p_{\mathbf{Q}}(S) \equiv Pr(\mathbf{Q} \in S) \tag{2.37}$$

  where $Pr(\mathbf{Q} \in S)$ is the probability that $\mathbf{Q}$ belongs to the set S.

- *Belief-mass function*: Let Ψ be a RFS.Then the belief-mass function $\beta_{\Psi}(S)$ is defined as:

$$\beta_{\Psi}(S) \equiv Pr(\Psi \subseteq S) \tag{2.38}$$

  Later, in this chapters it will be seen that the multi-object probability densities can be obtained from the belief-mass functions.

2.4.2 Set Integral

([12],chapter 11)

Consider a set function $f(Q)$ of a RFS Q. Then it's set integral over a region S of $\mathcal{Q}_0$ is:

$$\int_S f(Q)\delta Q \equiv \sum_{m=0}^{\infty} \frac{1}{m!} \int_{\underbrace{S \times ... \times S}_{m}} f(\{\mathbf{q}_1, ..., \mathbf{q}_m\}) d\mathbf{q}_1 ... d\mathbf{q}_m \tag{2.39}$$

$$= f(\emptyset) + \int_S f(\{\mathbf{q}\}) d\mathbf{q} + \frac{1}{2!} \int_{S \times S} f(\{\mathbf{q}_1, \mathbf{q}_2\}) d\mathbf{q} + ... \tag{2.40}$$

Consider the function $f_m(\mathbf{q}_1, ..., \mathbf{q}_m)$ in m vector variables $\mathbf{q}_1, ..., \mathbf{q}_m$.

$$f_m(\mathbf{q}_1, ..., \mathbf{q}_m) \equiv \begin{cases} \frac{1}{m!} f(\{\mathbf{q}_1, ..., \mathbf{q}_m\}) & \text{if } \mathbf{q}_1, ..., \mathbf{q}_m \text{ are distinct} \\ 0 & \text{if otherwise} \end{cases} \tag{2.41}$$

The division by $m!$ occurs because the elements of a RFS are unordered. Hence the probability density of $\{\mathbf{q}_1, ..., \mathbf{q}_m\}$ must be equally distributed among the $m!$ possible vectors $(\mathbf{q}_{\sigma_1}, ..., \mathbf{q}_{\sigma_m})$ for all permutations σ of $1, ..., m$. Then each term in equation 2.39 is equivalent to:

$$\int_{\underbrace{S \times ... \times S}_{m}} f(\{\mathbf{q}_1, ..., \mathbf{q}_m\}) d\mathbf{q}_1 ... d\mathbf{q}_m = m! \int_{\underbrace{S \times ... \times S}_{m}} f_m(\mathbf{q}_1, ..., \mathbf{q}_m) d\mathbf{q}_1 ... d\mathbf{q}_m \tag{2.42}$$

2.4.3 Multi-object Probability Distributions

The definition of *multi-object probability density function* is similar to probability density function of of a random variable. The only difference is that here the function is defined on sets. A *multi-object density function* $f(Q)$ on $\mathcal{Q}_0$ is a real-valued function of a finite set variable $Q \subseteq \mathcal{Q}_0$. A multi-object density function is a multi-object probability density function if $f(Q) \geq 0$ for all Q and if:

$$\int_{\mathcal{Q}_0} f(Q) \delta Q = 1 \tag{2.43}$$

The Belief-mass function can then be defined in terms of the multi-object probability density function by:

$$\beta_\Psi(S) = Pr(\Psi \subseteq S) = \int_S f_\Psi(Q) \delta Q \tag{2.44}$$

where Ψ is a RFS on $\mathcal{Q}_0$. Also if $S = \mathcal{Q}_0$, we get $\int_S f_\Psi(Q) \delta Q = Pr(\Psi \subseteq \mathcal{Q}_0) = 1$

2.4.4 Cardinality Distribution

The probability that a the RFS Ψ contains n elements is described by the cardinality distribution.

$$p_{\Psi}(m) = Pr(|\Psi| = m) \equiv p_{|\Psi|}(m) \tag{2.45}$$

$$= \int_{\underbrace{S \times ... \times S}_{m}} f(\{\mathbf{q}_1, ..., \mathbf{q}_m\}) d\mathbf{q}_1 ... d\mathbf{q}_m \tag{2.46}$$

where $S = \mathcal{Q}_0$

2.5 Probability-Generating Functionals(p.g.fls)

A functional is a real-valued function of a function i.e., it maps functions to real numbers. p.g.fls, as the name suggests, are functionals defined to reduce the complexity of the equations in multi-target filtering. This will become much clearer in the following chapters once the multi-target bayes filter equations are derived.
Consider a function $h(\mathbf{y})$, which is a non-negative real-valued function of $\mathbf{y} \in \mathfrak{Z}_0$. This is known as a test function. One important example of text function is the indicator function $\mathbb{1}_S(\mathbf{y})$, where $S \subseteq \mathfrak{Z}_0$ and is defined as:

$$\mathbb{1}_S(\mathbf{y}) = \begin{cases} 1 & \text{if } \mathbf{y} \in S \\ 0 & \text{if otherwise} \end{cases} \tag{2.47}$$

A functional F "evaluated" at the function h can be represented as $F[h]$. A simple example of a functional is the linear functional which is defined as:

$$f[h] \equiv \int h(\mathbf{y}) f(\mathbf{y}) d\mathbf{y} \tag{2.48}$$

where $f(\mathbf{y})$ is a density function on $\mathfrak{Z}_0$. Then, the functional evaluated at $\mathbb{1}_S$ is $f[\mathbb{1}_S] = \int_S f(\mathbf{y}) d\mathbf{y}$. Let $Y \subseteq \mathfrak{Z}_0$ be a finite subset whose elements belong to the space $\mathfrak{Z}_0$. Then the power of the test function(h) w.r.t Y is defined as:

$$h^Y \equiv \begin{cases} 1 & \text{if } Y = \emptyset \\ \prod_{\mathbf{y} \in Y} h(\mathbf{y}) & \text{if otherwise} \end{cases} \tag{2.49}$$

Then, the *probability-generating functional* of a probability density function $f_\Psi(Y)$ of a RFS Ψ is defined as:

$$G_\Psi[h] \equiv \int h^Y f_\Psi(Y)\delta Y \tag{2.50}$$

$$= f_\Psi(\emptyset) + \sum_{n=1}^{\infty} \frac{1}{n!} \int \left(\prod_{i=1}^{n} h(\mathbf{y}_i) \right) f_\Psi(\{\mathbf{y}_1, ..., \mathbf{y}_n\}) d\mathbf{y}_1 ... d\mathbf{y}_n \tag{2.51}$$

$$= f_\Psi(\emptyset) + \sum_{n=1}^{\infty} \int \left(\prod_{i=1}^{n} h(\mathbf{y}_i) \right) f_\Psi(\mathbf{y}_1, ..., \mathbf{y}_n) d\mathbf{y}_1 ... d\mathbf{y}_n \tag{2.52}$$

$$\tag{2.53}$$

Some important properties of a p.g.fl are given below:

- $G_\Psi[0] = f_\Psi(\emptyset)$
- $G_\Psi[1] = \int f_\Psi(Y)\delta Y = 1$
- If h($\mathbf{y}$)=y_0 where y_0 is a constant non-negative real number for all $\mathbf{y}$, then,

$$G_\Psi[h] = \int y_0^Y f_\Psi(Y)\delta Y \tag{2.54}$$

$$= f_\Psi(\emptyset) + y_0 \int f_\psi(\{\mathbf{y}\}) d\mathbf{y} + \frac{y_0^2}{2!} \int f_\psi(\{\mathbf{y}_1, \mathbf{y}_2\}) d\mathbf{y}_1 d\mathbf{y}_2 + ... \tag{2.55}$$

$$= p_\Psi(0) + y p_\Psi(1) + y^2 p_\Psi(2) + ... \tag{2.56}$$

$$= \sum_{n=0}^{\infty} y^n p_\Psi(n) \tag{2.57}$$

  where $p_\Psi(n)$ is the cardinality distribution of Ψ as discussed in section 2.4.4 .

- Belief-mass functions, defined in section 2.4.1 can ne obtained from p.g.fls in the following manner.

$$\beta_\Psi(S) = \int_S f_\Psi(Y)\delta Y = \int \mathbb{1}_S^Y f_\Psi(Y)\delta Y = G_\Psi[\mathbb{1}_S] \tag{2.58}$$

Until now, some important properties and quantities required to describe a RFS have been reviewed. The next section dwells into defining the concept of functional derivatives and formally defines the notion of set derivative.

2.6 Derivative of a Set

The concepts described in this section can be found in a detailed manner in ([12],chapter 11).

The *gradient derivative*(also known as the *Frechét/directional* derivative) of a vector-valued function $\tau(\mathbf{x})$ in the direction $\mathbf{w}$ is given by:

$$\frac{\partial \tau}{\partial \mathbf{w}}(\mathbf{x}) = \lim_{\epsilon \to 0} \frac{\tau(\mathbf{x} + \epsilon \mathbf{w}) - \tau(\mathbf{x})}{\epsilon} \tag{2.59}$$

Gradient derivative is reviewed here because a functional derivative is a special case of gradient derivative and the set derivative is a special case of functional derivative.The follwoing section reviews the concept of functional derivative and is followed by the definition of set derivative.

2.6.1 Functional Derivative

A functional $F[h]$ is a function which maps a function to a real number.Similar to the gradient derivative one can think of a functional derivative of a functional $F[h]$ at a function h in the direction of a function g (denoted by $\frac{\delta F}{\delta g}[h]$). Specifically, consider $g = \delta_{\mathbf{y}}$ where δ is a Dirac delta function and let $F[h]$ be a functional on $\mathfrak{Z}_0$ and let $y \in \mathfrak{Z}_0$. Then the **functional derivative** of $F[h]$ at $\mathbf{y}$ is (refer [34]):

$$\frac{\delta F}{\delta \mathbf{y}}[h] \equiv \lim_{\epsilon \to 0} \frac{F[h + \epsilon \delta_{\mathbf{y}}] - F[h]}{\epsilon} \tag{2.60}$$

Now, given $\mathbf{y}_1, ..., \mathbf{y}_m \in \mathfrak{Z}_0$, higher order functional derivatives of $F[h]$ at $\mathbf{y}_1, ..., \mathbf{y}_m$ is evaluated recursively by :

$$\frac{\delta^m F}{\delta \mathbf{y}_m ... \delta \mathbf{y}_1}[h] \equiv \frac{\delta}{\delta \mathbf{y}_m} \left(\frac{\delta^{m-1} F}{\delta \mathbf{y}_{m-1} ... \delta \mathbf{y}_1} \right)[h] \tag{2.61}$$

Hence, if Y is a finite subset whose elements belong to $\mathfrak{Z}_0$ then the functional derivative of $F[h]$ w.r.t Y is:

$$\frac{\delta F}{\delta Y}[h] \equiv \begin{cases} F[h] & \text{if } Y = \emptyset \\ \frac{\delta^m F}{\delta \mathbf{y}_m ... \delta \mathbf{y}_1}[h] & \text{if } Y = \{\mathbf{y}_1, ..., \mathbf{y}_m\} \end{cases} \tag{2.62}$$

The derivative of a linear functional(defined in equation 2.48) can be then given by:

$$\frac{\delta f}{\delta Y}[h] = \begin{cases} f[h] & \text{if } Y = \emptyset \\ f(\mathbf{y}) & \text{if } Y = \{\mathbf{y}\} \\ 0 & \text{if } |Y| \geq 2 \end{cases} \tag{2.63}$$

A detailed proof can be found in [12], page 378

2.6.2 Set Derivatives

As mentioned above, a set derivative can be thought of as a special case of a functional derivative.Every functional can give rise to a set function. For example, consider the functional $F[h]$. Then a set function $\phi(S)$ can be defined as:

$$\phi_F(S) \equiv F(\mathbb{1}_S) \tag{2.64}$$

Then the set derivative of the set function $\phi_F(S)$ is defined as(refer [12], section 11.4.2):

$$\frac{\delta \phi_F}{\delta Y}(S) \equiv \frac{\delta F}{\delta Y}[\mathbb{1}_S] \tag{2.65}$$

where the functional derivative $\frac{\delta F}{\delta Y}$ is defined as in equation 2.62. Applying equation 2.60, the functional derivative $\frac{\delta F}{\delta y}[\mathbb{1}_S]$ is:

$$\frac{\delta F}{\delta \mathbf{y}}[\mathbb{1}_S] = \lim_{\epsilon \to 0} \frac{F[\mathbb{1}_S + \epsilon \delta_y] - F[\mathbb{1}_S]}{\epsilon} \tag{2.66}$$

$$= \lim_{|D_\mathbf{y}| \to 0} \frac{F[\mathbb{1}_S + \mathbb{1}_{D_\mathbf{y}}] - F[\mathbb{1}_S]}{|D_\mathbf{y}|} \tag{2.67}$$

where $D_\mathbf{y}$ is a infinitesimal neighbourhood of $\mathbf{y}$ with the hypervolume $|D_y| = \epsilon$.Assume that S and $D_\mathbf{y}$ are disjoint.Then:

$$\frac{\delta F}{\delta \mathbf{y}}[\mathbb{1}_S] = \lim_{|D_\mathbf{y}| \to 0} \frac{F[\mathbb{1}_{S \cup D_\mathbf{y}}] - F[\mathbb{1}_S]}{|D_\mathbf{y}|} \tag{2.68}$$

$$= \lim_{|D_\mathbf{y}| \to 0} \frac{\phi_F(S \cup D_\mathbf{y}) - \phi_F(S)}{|D_\mathbf{y}|} \tag{2.69}$$

Therefore, the derivative of a set function $phi(S)$ can be formally defined as:

$$\frac{\delta\phi}{\delta\mathbf{y}}(S) \equiv \lim_{|D_{\mathbf{y}}|\to 0} \frac{\phi(S \cup D_{\mathbf{y}}) - \phi(S)}{|D_{\mathbf{y}}|} \tag{2.70}$$

Consequently the set derivative w.r.t to a finite set Y is given by:

$$\frac{\delta\phi}{\delta Y}(S) \equiv \begin{cases} \phi(S) & \text{if } Y = \emptyset \\ \frac{\delta^m \phi}{\delta\mathbf{y}_m \ldots \delta\mathbf{y}_1}(S) & \text{if } Y = \{\mathbf{y}_1, ..., \mathbf{y}_m\} \end{cases} \tag{2.71}$$

Now that the notion of a set derivative is established, the next section reviews some important theorems in multi-object calculus and states the fundamental rules for set differentiation analogous to the rules like the chain rule in vector calculus.

2.6.3 Radon-Nikodým Theorems

The fundamental theorem of multi-object calculus states that the set integral and set derivative are inverse operations(refer [12], [35]).It is stated as follows:

$$\phi(S) = \int_S \frac{\delta\phi}{\delta Y}(\emptyset)\delta Y \tag{2.72}$$

$$\left[\frac{\delta}{\delta Y}\int_S f(W)\delta W\right]_{S=\emptyset} = f(Y) \tag{2.73}$$

The fundamental theorem applied on the belief-mass function(a set function as defined in equation 2.38) gives rise to the first Radon-Nikodým theorem which is stated as (refer [12], [35]):

$$\int_S \frac{\delta\beta_\Psi}{\delta Y}(\emptyset)\delta Y = Pr(\Psi \subseteq S) = \beta_\Psi(S) \tag{2.74}$$

The multi-object probability density function is related to the belief-mass function by equation 2.44. Using this fact and the Radon-Nikodým theorem it can be seen that the probability density function for a RFS Ψ can be obtained from the set derivatives of the belief-mass function as:

$$\beta_\Psi(S) = \int_S f_\Psi(Y)\delta Y, \text{ which implies, } f_\Psi(Y) = \frac{\delta\beta_\Psi}{\delta Y}(\emptyset) \tag{2.75}$$

Equation 2.75 is central to the development of the multi-target filtering equations.The equation 2.75 can further be modified to express the set derivatives of $\beta_\Psi(S)$ as follows(refer [36]):

$$\frac{\delta\beta_\Psi}{\delta Y}(S) = \int_S f_\Psi(Y \cup W)\delta W \tag{2.76}$$

It can be further generalised to p.g.fls as:

$$\frac{\delta G_\Psi}{\delta Y}[h] = \int h^W f_\Psi(Y \cup W)\delta W \tag{2.77}$$

The probability density function of a RFS can also be obtained as:

$$f_\Psi(Y) = \frac{\delta G_\Psi}{\delta Y}[0] \tag{2.78}$$

2.6.4 Basic Differentiation Rules

The following rules(refer [12](section 11.6) and [35]) are fundamental to the derivation of the multi-target filtering equations.They are:

- *Constant rule*: Let $\phi(S) = K$ be a constant set function.Then the rule is stated as:

$$\frac{\delta}{\delta Y}K = 0 \tag{2.79}$$

- *Linear rule*:Consider the set function($p_f(S)$) defined by $p_f(S) = \int_S f(\mathbf{y})d\mathbf{y}$ for all S.Then:

$$\frac{\delta}{\delta Y}p_f(S) = \begin{cases} p_f(S) & \text{if } Y = \emptyset \\ f(\mathbf{y}) & \text{if } Y = \{y\} \\ 0 & \text{if } |Y| \geq 2 \end{cases} \tag{2.80}$$

- *Sum rule*:$\phi_1(S), \phi_2(S)$ are two set functions and a_1, a_2 are two real numbers.Then the sum rule is:

$$\frac{\delta}{\delta Y}(a_1\phi_1(S) + a_2\phi_2(S)) = a_1\frac{\delta\phi_1}{\delta Y}(S) + a_2\frac{\delta\phi_2}{\delta Y}(S) \tag{2.81}$$

this can be extended to a sum of any number of set functions.

- *Product Rule*:Let $\phi_1, ..., \phi_m$ be m set functions.Then:

$$\frac{\delta}{\delta Y}(\phi_1(S)...\phi_m(S)) = \sum_{W_1 \cup ... \cup W_m = Y} \frac{\delta \phi_1}{\delta W_1}(S)...\frac{\delta \phi_m}{\delta W_m}(S) \tag{2.82}$$

  where the summation is taken over all the mutually disjoint subsets $W_1, ..., W_m$ of Y whose union is Y.

- *Chain rule*: Let $f(y_1, ..., y_m)$ be real-valued functions of the real variables $y_1, ..., y_m$ and $\phi_1, ..., \phi_m$ be m set funcitons.

$$\frac{\delta}{\delta \mathbf{y}} f(\phi_1(S), ..., \phi_m(S)) = \sum_{i=1}^{m} \frac{\delta f}{\delta \mathbf{y}_i}(\phi_1(S), ..., \phi_m(S))\frac{\delta \phi_i}{\delta \mathbf{y}}(S) \tag{2.83}$$

- All the above rules are just special cases of rules of functional derivatives. For example, the sum rule for functional derivatives is given as follows:

$$\frac{\delta}{\delta Y}(a_1 F_1[h] + a_2 F_2[h]) = a_1 \frac{\delta F}{\delta Y}[h] + a_2 \frac{\delta F_2}{\delta Y}[h] \tag{2.84}$$

  where F_1, F_2 are functionals and a_2, a_2 are real numbers.All the other rules(constant,product,chain) can be given in a similar fashion.

2.7 Independent Random Finite Sets

The concept of independent RFS is widely used in multi-target filtering. In the modeling of motion of targets, it is often assumed that targets move independent of each other. In the case of measurements it is often assumed that the targets generate measurements independently and also that the clutter is generated independently. How these concepts manifest in terms of RFS will be seen in the following chapter.Consider a RFS $\Psi = \Psi_1 \cup ... \cup \Psi_n$ where $\Psi_1, ..., \Psi_n$ are statistically independent random finite subsets.The probability density of Ψ can be then given by:

$$f_\Psi(Y) = \sum_{W_1 \cup ... \cup W_n = Y} f_{\Psi_1}(W_1)...f_{\Psi_n}(W_n) \tag{2.85}$$

where the summation is taken over all the mutually disjoint subsets $W_1, ..., W_n$ of Y whose union is Y.This can be obtained by the direct application of the product rule(equation 2.82) on the belief-mass functions,i.e.,:

$$\frac{\delta\beta_{\Psi}}{\delta Y}(S) = \sum_{W_1 \cup ... \cup W_n} \frac{\delta\beta_{\Psi_1}}{\delta W_1}(S)...\frac{\delta\beta_{\Psi_n}}{\delta W_n}(S) \tag{2.86}$$

Then substituting $S = \emptyset$ and using the equation 2.75, the probability density function f_{Ψ} can be given by:

$$f_{\psi}(Y) = \frac{\delta\beta_{\Psi}}{\delta Y}(\emptyset) \tag{2.87}$$

$$= \sum_{W_1 \cup ... \cup W_n} \frac{\delta\beta_{\Psi_1}}{\delta W_1}(\emptyset)...\frac{\delta\beta_{\Psi_n}}{\delta W_n}(\emptyset) \tag{2.88}$$

$$= \sum_{W_1 \cup ... \cup W_n} f_{\Psi_1}(W_1)...f_{\Psi_n}(W_n) \tag{2.89}$$

The p.g.fl of Ψ is given by:

$$G_{\Psi_1 \cup ... \cup \Psi_n}[h] = G_{\Psi_1}[h]...G_{\Psi_n}[h] \tag{2.90}$$

This can be derived as follows:

$$G_{\Psi_1 \cup ... \cup \Psi_n}[h] = \int h^Y . f_{\Psi_1 \cup ... \cup \Psi_n}(Y)\delta Y \tag{2.91}$$

$$= \int h^{W_1 \cup ... \cup W_n} \left(\sum_{W_1 \cup ... \cup W_n = Y} f_{\Psi_1}(W_1)...f_{\Psi_n}(W_n) \right) \delta Y \tag{2.92}$$

$$= \int h^{W_1}...h^{W_n} . f_{\Psi_1}(W_1)...f_{\Psi_n}(W_n)\delta W_1...\delta W_n \tag{2.93}$$

The last equation is a consequence of equation 2.49 and the fact that the sets $W_1, ..., W_n$ are mutually disjoint.Thus,

$$G_{\Psi_1 \cup ... \cup \Psi_n}[h] = \left(\int h^{W_1} . f_{\Psi_1}(W_1)\delta W_1 \right) ... \left(\int h^{W_n} . f_{\Psi_n}(W_n)\delta W_n \right) \tag{2.94}$$

$$= G_{\Psi_1}[h]...G_{\Psi_n}[h] \tag{2.95}$$

2.8 Examples of Random Finite Sets and their p.g.fls

This section introduces two important class of RFS which are often used in multi-object filtering.

2.8.1 Bernoulli RFS

A Bernoulli RFS is described by a parameter p where $0 \leq p \leq 1$.Consider a RFS Ψ such that:

$$|\Psi| = \begin{cases} 0 & \text{with probability p} \\ 1 & \text{with probability 1-p} \end{cases} \tag{2.96}$$

where $|\Psi|$ is the cardinality of Ψ.Then Ψ can be expressed as:

$$\Psi = \emptyset^p \cap \{\mathbf{Y}\} \tag{2.97}$$

where $\mathbf{Y}$ is random vector belonging to a vector space $\mathfrak{Z}_0$ and is defined by:

$$Pr(\emptyset^p = S) \equiv \begin{cases} 1-p & \text{if S=}\emptyset \\ p & \text{if S=}\mathfrak{Z}_0 \\ 0 & \text{otherwise} \end{cases} \tag{2.98}$$

The belief-mass function (equation 2.38) is then given by([12],page 351):

$$\beta_\Psi(S) = Pr(\Psi \subseteq S) \tag{2.99}$$

$$= Pr(\Psi \subseteq S, \Psi \neq \emptyset) + Pr(\Psi \subseteq S, \Psi = \emptyset) \tag{2.100}$$

$$= Pr(\Psi \neq \emptyset).Pr(\Psi \subseteq S|\Psi \neq \emptyset) + Pr(\Psi = \emptyset) \tag{2.101}$$

$$= p.Pr(Y \in S|\Psi \neq \emptyset) + 1 - p \tag{2.102}$$

$$= 1 - p + p \int_S f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} \tag{2.103}$$

where $f_{\mathbf{Y}}(\mathbf{y})$ is the probability density function(pdf) of the random vector $\mathbf{Y}$.The probability density function of the RFS Ψ is defined as:

$$f_\psi(Y) \equiv \begin{cases} p.f_{\mathbf{Y}}(\mathbf{y}) & \text{if } Y = \{\mathbf{y}\} \\ 1-p & \text{if } Y = \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{2.104}$$

It is also seen that the relation between belief-mass function and the probability density function(equation 2.44) is verified.Detailed proof can be found in [12](page 352).The p.g.fl of a Bernoulli RFS is can be derived as follows:

$$G_{\Psi}[h] = \int h^{Y} f_{\Psi}(Y)\delta Y \tag{2.105}$$

$$= f_{\Psi}(\emptyset) + \int h(\mathbf{y}) f_{\Psi}(\{\mathbf{y}\}) d\mathbf{y} \tag{2.106}$$

$$= 1 - p + p \int h(\mathbf{y}) f_{\mathbf{Y}}(\mathbf{y}) d\mathbf{y} \tag{2.107}$$

As will be seen in the following chapters Bernoulli RFS is an essential component in multi-object filtering since it used to model the motion of a single object and also describe the behaviour of a single measurement.

2.8.2 Poisson RFS

The Poisson RFS is analogous to the Poisson random vector.The number of elements in this RFS is described by a Poisson distribution and each element is independently identically distributed(i.i.d) i.e. if a $\{\mathbf{y}_1, ..., \mathbf{y}_n\}$ is a finite set, then $y_1, ..., y_n$ are independent to each other and have the same distribution $f(\mathbf{y})$.The cardinality distribution of a Poisson RFS Ψ is then given by:

$$p_{|\Psi|}(n) = \frac{e^{-\lambda}\lambda^{n}}{n!} \tag{2.108}$$

where $\lambda \geq 0$ is a parameter(also the mean of $p_{|\Psi|}(n)$).Let the probability density function of the elements be $f(\mathbf{y})$ where $\mathbf{y} \in \mathfrak{Z}_0$. Then the probability distribution function of Ψ is given by:

$$f_{\Psi}(Y) \equiv n!.p_{|\Psi|}(n).f(\mathbf{y}_1)...f(\mathbf{y}_n) \tag{2.109}$$

$$= e^{-\lambda}.\lambda^{n}.f(\mathbf{y}_1)...f(\mathbf{y}_n) \tag{2.110}$$

where $Y = \{\mathbf{y}_1, ..., \mathbf{y}_n\}$ with $|Y| = n$. It is easy to verify that $f_\Psi(Y)$ is a probability distribution by checking the condition $\int f_\Psi(Y)\delta Y = 1$. Proof can be found in [12](page 799). The belief mass function(equation 2.38) is then given by:

$$\beta_\Psi(S) = \int_S f_\Psi(Y)\delta Y \tag{2.111}$$

$$= f_\Psi(\emptyset) + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{S\times...\times S} f(\{\mathbf{y}_1, ..., \mathbf{y}_n\}) d\mathbf{y}_1 ... d\mathbf{y}_n \tag{2.112}$$

$$= e^{-\lambda} + e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \int_{S\times...\times S} f(\mathbf{y}_1)...f(\mathbf{y}_n) d\mathbf{y}_1 ... d\mathbf{y}_n \tag{2.113}$$

$$= e^{-\lambda} + e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \left(\int_S f(\mathbf{y}) d\mathbf{y} \right)^n \tag{2.114}$$

$$= e^{\lambda \int_S f(\mathbf{y}) d\mathbf{y} - \lambda} = e^{\lambda(\int_S f(\mathbf{y}) d\mathbf{y} - 1)} \tag{2.115}$$

where the last equation is obtained by using the Taylor series expansion of a exponential. The p.g.fl of the Poisson RFS Ψ can be derived as follows:

$$G_\Psi[h] = \int h^Y f_\Psi(Y)\delta Y \tag{2.116}$$

$$= \sum_{n\geq 0} \int \left(\prod_{i=1}^{n} h(\mathbf{y}_i) \right) \frac{e^{-\lambda}\lambda^n}{n!} \left(\prod_{i=1}^{n} f(\mathbf{y}_i) \right) d\mathbf{y}_1 ... d\mathbf{y}_n \tag{2.117}$$

$$= e^{-\lambda} \sum_{n\geq 0} \frac{\lambda^n}{n!} \int \left(\prod_{i=1}^{n} h(\mathbf{y}_i) f(\mathbf{y}_i) \right) d\mathbf{y}_1 ... d\mathbf{y}_n \tag{2.118}$$

$$= e^{-\lambda} \sum_{n\geq 0} \frac{\lambda^n}{n!} \int \left(h(\mathbf{y}) f(\mathbf{y}) d\mathbf{y} \right)^n \tag{2.119}$$

$$= e^{-\lambda} \sum_{n\geq 0} \frac{(\lambda \int h(\mathbf{y}) f(\mathbf{y}) d\mathbf{y})^n}{n!} \tag{2.120}$$

$$= e^{\lambda(\int h(\mathbf{y}) f(\mathbf{y}) d\mathbf{y} - 1)} \tag{2.121}$$

where once again the last equation is obtained by using the Taylor series expansion of e^x.

This chapter has introduced the toolbox necessary to develop a systematic approach to derive the multi-target filtering equations. Hence, in the next chapter the multi-target Bayes filter will be formally derived.

3. THE MULTI-TARGET BAYES FILTER

Similar to the case of a single target filter, the first step is to model the dynamics of the multiple targets and the measurements.The next step is to formulate them in terms of random finite sets and use FISST (Finite Set Statistics) to construct the multi-target Markov and likelihood densities. Finally, the Bayes rule is used to obtain a recursive Multi-target Bayes Filter analogous to the single target case.
The following section models the multi-target measurements and methodically derives the probability densities required to describe the multi-target measurements. The derivation given in this chapter was obtained from [12].

3.1 Multi-Target Likelihood Functions

This section deals with the derivation of likelihood function for multi-target measurements. The following steps detail the procedure to find the likelihood function.

- The first step is to define the RFS $\Sigma_{k+1}(X)$, describing the randomly varying measurement set collected by a single sensor at time step $k+1$ given that the targets have state X.

- The belief mass function (equation 2.38) corresponding to the RFS $\Sigma_{k+1}(X)$ completely characterizes statistics of the multi-target measurement set.

$$\beta_{k+1}(T|X) = Pr(\Sigma_{k+1} \subseteq T|X) \tag{3.1}$$

- The probability density function of Σ_{k+1} i.e., $f_{k+1}(Z|X)$ is then obtained from belief mass function (equation 2.75) as follows:

$$f_{k+1}(Z|X) = \frac{\delta\beta_{k+1}}{\delta Z}(\emptyset|X) \tag{3.2}$$

$$= \left[\frac{\delta\beta_{k+1}}{\delta Z}(T|X)\right]_{T=\emptyset} \tag{3.3}$$

The probability density $f_{k+1}(Z|X)$ is the *true multi-target likelihood function.*

3.1.1 Multi-target Measurement Space

The multi-target measurement space is identical to the example given in Section 2.4. As was mentioned before, a *multi-target measurement* is best represented by a Random Finite Set. Let Z be a RFS representing the multi-target measurement, then:

$$Z = \{\mathbf{z}_1, ..., \mathbf{z}_m\} \tag{3.4}$$

is a instantiation (a sample) of the RFS where $\mathbf{z}_1, ..., \mathbf{z}_m$ are elements of Z which belong to a *single-sensor measurement space* $\mathfrak{Z}_0$. Then the various instantiations of Z can be given by:

$$Z = \emptyset \text{ (no detections)} \tag{3.5}$$

$$Z = \{\mathbf{z}_1\} \text{ (a single detection } \mathbf{z}_1 \text{ is collected)} \tag{3.6}$$

$$Z = \{\mathbf{z}_1, \mathbf{z}_2\} \text{ (two detections } \mathbf{z}_1 \neq \mathbf{z}_2 \text{ are collected)} \tag{3.7}$$

and a sample with m detections is given by:

$$Z = \{\mathbf{z}_1, ..., \mathbf{z}_m\} \text{ (m detections } \mathbf{z}_1 \neq ... \neq \mathbf{z}_m \text{ are collected)} \tag{3.8}$$

Then the *single-sensor multi-target measurement space* $\mathfrak{Z}$ is the hyperspace of all the finite sets whose elements belong to the single-sensor measurement space $\mathfrak{Z}_0$.

3.1.2 Standard Measurement Model

The measurement set Z can be divided into two parts where one part consists of measurements from the targets and the other part consists of the clutter measurements. There are also certain assumptions to be made in order to derive the likelihood function $f_{k+1}(Z|X)$. They are:

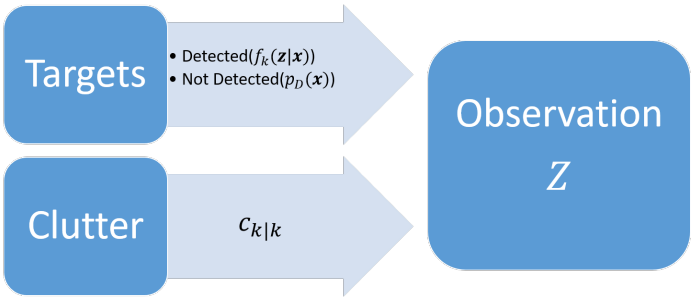


Figure 3.1. Standard Measurement Model

- A single sensor with state $\tilde{\mathbf{x}}$ observes a scene with unknown number of targets with the single target likelihood function:

$$f_{k+1}(\mathbf{z}|\mathbf{x},\tilde{\mathbf{x}}) \equiv f_{k+1}(\mathbf{z}|\mathbf{x}) = f_{\mathbf{W}_{k+1}}(\mathbf{z} - \eta_{k+1}(\mathbf{x},\tilde{\mathbf{x}})) \tag{3.9}$$

  where $\eta_{k+1}(\mathbf{x},\tilde{\mathbf{x}}) + \mathbf{W}_{k+1}(\mathbf{W}_{k+1}$ is measurement noise) is the single-target measurement model i.e., measurement obtained from a target with state $\mathbf{x}$ and sensor state $\tilde{\mathbf{x}}$.

- All the measurements are unique i.e., no measurement is generated by more than one target.

- A single target with a state $\mathbf{x}$ in the scene, either generates a measurement with the probability $p_D(\mathbf{x}) \equiv p_D(\tilde{\mathbf{x}})$ or is undetected i.e., generates no measurement with the probability $1 - p_D(\mathbf{x})$.It is assumed that individual target measurement sets are Bernoulli RFS(section 2.8.1).

- The false measurements i.e., the clutter measurement is modeled as a Poisson RFS $C \equiv C(\tilde{\mathbf{x}})$(Section 2.8.2) with a parameter $\lambda \equiv \lambda(\tilde{\mathbf{x}})$ and the elements are distributed in space according to the density $c(\mathbf{z}) \equiv c(\mathbf{z}|\tilde{\mathbf{x}})$.

- It is also assumed that the clutter measurements(C) and the target generated measurements($\Upsilon(X)$) are statistically independent.

3.1.3 Measurement RFS

Let $X = \mathbf{x}_1, ..., \mathbf{x}_n$ be a predicted multi-target state at time step t_{k+1}. Then, the measurement set(Σ_{k+1}) collected by the sensor at t_{k+1} will be of the form:

$$\overset{\text{total measurement set}}{\Sigma_{k+1}} = \overset{\text{target detection set}}{\Upsilon(X)} \cup \overset{\text{clutter measurement set}}{C} \tag{3.10}$$

i.e., if Z is an instantiation of Σ_{k+1},then:

$$\overset{\text{total measurement set}}{Z} = \overset{\text{target detection set}}{\{\mathbf{z}_1, ..., \mathbf{z}_{n'}\}} \cup \overset{\text{clutter measurement set}}{\{\mathbf{c}_1, ..., \mathbf{c}_{m'}\}} \tag{3.11}$$

Furthermore, the target measurement set $\Upsilon(X)$ can be expressed as a union of n RFS as follows:

$$\overset{\text{target detection set}}{\Upsilon(X)} = \overset{\text{detection set from } \mathbf{x}_1}{\Upsilon(\mathbf{x}_1)} \cup ... \cup \overset{\text{detection set from } \mathbf{x}_n}{\Upsilon(\mathbf{x}_n)} \tag{3.12}$$

As mentioned in the assumptions, $\Upsilon(\mathbf{x}_i), i = 1, ..., n$ are modeled as Bernoulli RFS and have the form:

$$\Upsilon(\mathbf{x}_i) = \emptyset^{p_D(\mathbf{x}_i)} \cap \mathbf{Z}_i \tag{3.13}$$

where $\mathbf{Z}_i = \eta(\mathbf{x}_i), \mathbf{Z} \in \mathfrak{Z}_0$ is the sensor-measurement model and $\emptyset^{p_D(\mathbf{x}_i)}$ is a discrete RFS (section 2.8.1) defined as follows.

$$Pr(\emptyset^{p_D(\mathbf{x}_i)} = T) = \begin{cases} 1 - p_D(\mathbf{x}_i) & \text{if } T = \emptyset \\ p_D(\mathbf{x}_i) & \text{if } T = \mathfrak{Z}_0 \\ 0 & \text{if otherwise} \end{cases} \tag{3.14}$$

The probability density function of $\Upsilon(\mathbf{x}_i)$ is then by equation 2.104 as:

$$f_{\Upsilon(\mathbf{x}_i)}(Z) = \begin{cases} p_D(\mathbf{x}_i).f_{k+1}(\mathbf{z}|\mathbf{x}_i) & \text{if } Z = \{\mathbf{z}\} \\ 1 - p_D(\mathbf{x}_i) & \text{if } Z = \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{3.15}$$

The belief-mass function of $\Upsilon(\mathbf{x}_i)$ is then given by (equation 2.103):

$$\beta_{\Upsilon(\mathbf{x}_i)}(S) = 1 - p_D(\mathbf{x}_i) + p_D(\mathbf{x}_i) \int_S f_{k+1}(\mathbf{z}|\mathbf{x}_i) d\mathbf{z} \tag{3.16}$$

As mentioned before, it is assumed that C is a Poisson RFS(section 2.8.2) with the parameter λ and spatial distribution $c(\mathbf{z}) = f_C(\mathbf{z})$.Then if $C = \{\mathbf{c}_1, ..., \mathbf{c}_{m'}\}$, the cardinality distribution of C($|C| = m', m' \geq 0$) is:

$$p_{|C|}(m') = \frac{e^{-\lambda}\lambda^{m'}}{m'!}$$

and conditioned on $|C| = m'$,the elements $\mathbf{c}_1, ..., \mathbf{c}_{m'}$ are independent and identically distributed (i.i.d) random vectors with probability density $c(\mathbf{z}) = f_C(\mathbf{z})$.
Finally it is also assumed that $\Upsilon(\mathbf{x}_1), ..., \Upsilon(\mathbf{x}_n), C$ are statistically independent RFS.

3.1.4 Derivation of likelihood function($f_{k+1}(Z|X)$)

The derivation given here is obtained from [12], Appendix G.18

Consider the belief-mass function (defined as in equation 2.38) of $\Sigma_{k+1}(T|X)$ where $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$. Then:

$$\begin{aligned}
\beta_{k+1}(T|X) &= Pr(\Upsilon(X) \cup C \subseteq T|X) && (3.17)\\
&= Pr(\Upsilon(X) \subseteq T|X, C \subseteq T|X) && (3.18)\\
&= Pr(\Upsilon(X) \subseteq T|X).Pr(C \subseteq T) && (3.19)\\
&= \beta_{\Upsilon(X)}(T).\beta_C(T) && (3.20)
\end{aligned}$$

The assumption of independent RFS was used to obtain the above equations. Then, using the product rule (equation 2.82), the derivative of β_{k+1} is obtained as:

$$\frac{\delta\beta_{k+1}}{\delta Z}(T|X) = \sum_{W \subseteq Z} \frac{\delta\beta_{\Upsilon(X)}}{\delta W}(T).\frac{\delta\beta_C}{\delta Z}(T) \quad (3.21)$$

Using equation 2.75 i.e.,setting $T = \emptyset$ we obtain the likelihood function $f_{k+1}(Z|X)$ as:

$$\begin{aligned}
f_{k+1}(Z|X) &= \frac{\delta\beta_{k+1}}{\delta Z}(\emptyset|X) && (3.22)\\
&= \sum_{W \subseteq Z} f_{\Upsilon(X)}(W).f_C(Z - W) && (3.23)
\end{aligned}$$

where $f_{\Upsilon(X)}(Z)$ and $f_C(Z)$ are the probability density functions of $\Upsilon(X)$ and C respectively.

Probability Density of Clutter RFS($f_C(Z)$)

Since C is a Poisson RFS, it's belief-mass function is given by (equation 2.115):

$$\beta_C(T) = Pr(C \subseteq T) = e^{\lambda p_c(T) - \lambda} \tag{3.24}$$

where $p_c(T) = \int_T c(\mathbf{z}) d\mathbf{z}$. Then from equation 2.71, applying the set derivative on $\beta_C(T)$:

$$\frac{\delta \beta_C}{\delta Z}(T) = \frac{\delta}{\delta Z} e^{\lambda p_c(T) - \lambda} = \frac{\delta^m}{\delta \mathbf{z}_m ... \delta \mathbf{z}_1} e^{\lambda p_c(T) - \lambda} \tag{3.25}$$

$$= \frac{\delta^{m-1}}{\delta \mathbf{z}_m ... \delta \mathbf{z}_2} \frac{\delta}{\delta \mathbf{z}_1} e^{\lambda p_c(T) - \lambda} \tag{3.26}$$

Using the chain rule (equation 2.83):

$$\frac{\delta \beta_C}{\delta Z}(T) = \frac{\delta^{m-1}}{\delta \mathbf{z}_m ... \delta \mathbf{z}_2} \frac{\delta}{\delta \mathbf{z}_1} e^{\lambda p_c(T) - \lambda} = \frac{\delta^{m-1}}{\delta \mathbf{z}_m ... \delta \mathbf{z}_2} e^{\lambda p_c(T) - \lambda} \frac{\delta}{\delta \mathbf{z}_1} (\lambda p_c(T) - \lambda) \tag{3.27}$$

Applying the linear rule (equation 2.80):

$$\frac{\delta}{\delta \mathbf{z}_i} (\lambda p_c(T) - \lambda) = \lambda . c(\mathbf{z}_i); i = 1, ..., m \tag{3.28}$$

Then,

$$\frac{\delta \beta_C}{\delta Z}(T) = \frac{\delta^{m-1}}{\delta \mathbf{z}_m ... \delta \mathbf{z}_2} e^{\lambda p_c(T) - \lambda} \frac{\delta}{\delta \mathbf{z}_1} (\lambda p_c(T) - \lambda) \tag{3.29}$$

$$= \lambda . c(\mathbf{z}_1) . \frac{\delta^{m-2}}{\delta \mathbf{z}_m ... \delta \mathbf{z}_3} \frac{\delta}{\delta \mathbf{z}_2} e^{\lambda p_c(T) - \lambda} \tag{3.30}$$

$$= \lambda^2 . c(\mathbf{z}_1) . c(\mathbf{z}_2) \frac{\delta^{m-2}}{\delta \mathbf{z}_m ... \delta \mathbf{z}_3} e^{\lambda p_c(T) - \lambda} \tag{3.31}$$

$$= ... = \lambda^m . c(z_1) ... c(z_m) . e^{\lambda p_c(T) - \lambda} \tag{3.32}$$

$f_C(T)$ is given by (equation 2.75), i.e.:

$$f_C(Z) = \frac{\delta \beta_C}{\delta Z}(\emptyset) \tag{3.33}$$

$$= e^{-\lambda} \lambda^m c(\mathbf{z}_1) ... c(\mathbf{z}_m) \tag{3.34}$$

Probability Density of target generated measurement RFS($f_{\Upsilon(X)}(Z)$)

Consider,

$$\overset{\text{target detection set}}{\Upsilon(X)} = \overset{\text{detection set from } \mathbf{x}_1}{\Upsilon(\mathbf{x}_1)} \cup ... \cup \overset{\text{detection set from } \mathbf{x}_n}{\Upsilon(\mathbf{x}_n)} \tag{3.35}$$

where $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$.Since $\Upsilon(\mathbf{x}_1), ..., \Upsilon(\mathbf{x}_n)$ are independent RFS (assumption), the fundamental convolution formula (equation 2.88) is applied to obtain:

$$f_{\Upsilon(X)}(Z) = \frac{\delta \beta_{\Upsilon(X)}}{\delta Z}(\emptyset) = \sum_{W_1 \cup ... \cup W_n = Z} \frac{\delta p_1}{\delta W_1}(\emptyset) ... \frac{\delta p_n}{\delta W_n}(\emptyset) \tag{3.36}$$

where $p_i(S)$ is the belief-mass function of $\Upsilon(\mathbf{x}_i)$ (as given in equation 3.16), $Z = \{\mathbf{z}_1, ..., \mathbf{z}_m\}, m \leq n$ and $W_1, ..., W_n$ are mutually disjoint subsets of Z such that $W_1 \cup ... \cup W_n = Z$. Then, using the definition of set derivative (equation 2.71) and constant rule (equation 2.79):

$$\frac{\delta p_i}{\delta \emptyset}(\emptyset) = p_i(\emptyset) = 1 - p_D(\mathbf{x}_i) \tag{3.37}$$

$$\frac{\delta p_i}{\delta \{\mathbf{z}\}}(\emptyset) = p_D(\mathbf{x}_i).f_{k+1}(\mathbf{z}|\mathbf{x}_i) \tag{3.38}$$

$$\frac{\delta p_i}{\delta Z} = 0 \text{ if } |Z| > 1 \tag{3.39}$$

If $Z = \emptyset$ then the summation in equation 3.36 will involve only one term which has $W_1 = ... = W_n = \emptyset$. Then using equation 3.37,

$$f_{\Upsilon(X)}(\emptyset) = \frac{\delta p_1}{\delta \emptyset}(\emptyset) ... \frac{\delta p_n}{\delta \emptyset}(\emptyset) \tag{3.40}$$

$$= \prod_{i=1}^{n} (1 - p_D(\mathbf{x}_i)) \tag{3.41}$$

If $Z \neq \emptyset$ i.e. $Z = \{\mathbf{z}_1, ..., \mathbf{z}_m\}$,then:

$$f_{\Upsilon(X)}(Z) = f_{\Upsilon(X)}(\emptyset) \sum_{W_1 \cup ... \cup W_n = Z} \frac{\frac{\delta p_1}{\delta W_1}(\emptyset) ... \frac{\delta p_n}{\delta W_n}(\emptyset)}{\prod_{i=1}^{n} (1 - p_D(\mathbf{x}_i))} \tag{3.42}$$

$$= f_{\Upsilon(X)}(\emptyset) \sum_{1 \leq i_1 \neq ... \neq i_m \leq n} \frac{\frac{\delta p_{i_1}}{\delta \mathbf{z}_1}(\emptyset) ... \frac{\delta p_{i_m}}{\delta \mathbf{z}_m}(\emptyset)}{(1 - p_D(\mathbf{x}_{i_1})) ... (1 - p_D(\mathbf{x}_{i_m}))} \tag{3.43}$$

$$\tag{3.44}$$

Using equation 3.38:

$$f_{\Upsilon(X)}(Z) = f_{\Upsilon(X)}(\emptyset) \sum_{1 \leq i_1 \neq ... \neq i_m \leq n} \left(\prod_{j=1}^{m} \frac{p_D(\mathbf{x}_{i_j}).f_{k+1}(\mathbf{z}_j|\mathbf{x}_{i_j})}{(1 - p_D(\mathbf{x}_{i_j}))} \right) \tag{3.45}$$

The summation is now over the m-tuple $(i_1, ..., i_m)$ with $1 \leq i_1 \neq ... \neq i_m \leq n$ which defines a one-to-one function $\tau(j) = i_j$ where $\tau : \{1, ..., m\} \rightarrow \{1, ..., n\}$.For each one-to-one function, a inverse function θ can be defined as $\theta(i) = \tau^{-}(i)$ where $\theta : \{1, ..., n\} \rightarrow \{0, 1, ..., m\}$.

- For every $i = 1, ..., n$ if $\theta(i) = k, k > 0$ then it means that the state $\mathbf{x}_i$ uniquely generates the measurement $\mathbf{z}_k$ and if $\theta(i) = 0$ then it means that the target with state $\mathbf{x}_i$ is not detected.

- This basically ensures that the condition of $W_1 \cup ... \cup W_n = Y$ and that $W_1, ..., W_n$ are disjoint sets.

Therefore, for every m-tuple there exists an association θ as defined above and hence the summation can be rewritten in terms of θ.Therefore,for $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$,$f_{\Upsilon(X)}(Z)$ is:

$$f_{\Upsilon(X)}(Z) = \begin{cases} \prod_{i=1}^{n}(1 - p_D(\mathbf{x}_i)) & \text{if } Z = \emptyset \\ f_{\Upsilon(X)}(\emptyset) \sum_{\theta} \left(\prod_{i:\theta(i)>0} \frac{p_D(\mathbf{x}_i).f_{k+1}(\mathbf{z}_{\theta(i)}|\mathbf{x}_i)}{(1-p_D(\mathbf{x}_i))} \right) & \text{if } Z = \{\mathbf{z}_1, ..., \mathbf{z}_m\} \end{cases} \tag{3.46}$$

Using equation 3.34, it can be seen that:

$$f_C(Z - W) = e^{-\lambda} \prod_{\mathbf{z} \in Z - W} \lambda c(\mathbf{z}) \tag{3.47}$$

From equation 3.46 it can be seen that if $W = \{\mathbf{w}_1, ..., \mathbf{w}_e\}$, then $f_{\Upsilon(X)}(Z) = 0$ if $e > n$, otherwise it is seen that:

$$f_{\Upsilon(X)}(W) = f_{\Upsilon(X)}(\emptyset) \sum_{\theta} \left(\prod_{i:\theta(i)>0} \frac{p_D(\mathbf{x}_i).f_{k+1}(\mathbf{z}_{\theta(i)}|\mathbf{x}_i)}{(1 - p_D(\mathbf{x}_i))} \right) \tag{3.48}$$

Here $\theta : \{1, ..., n\} \rightarrow \{0, 1, ..., e\}$. Substituting the above terms in equation in 3.23, it is seen that:

$$f_{k+1}(Z|X) = \sum_{W \subseteq Z, |W| \leq n} f_{\Upsilon(X)}(W).f_C(Z - W) \tag{3.49}$$

$$= e^{-\lambda} f_{\Upsilon(X)}(\emptyset) \sum_{W \subseteq Z, |W| \leq n} \left(\sum_{\theta} \prod_{i:\theta(i)>0} \frac{p_D(\mathbf{x}_i).f_{k+1}(\mathbf{z}_{\theta(i)}|\mathbf{x}_i)}{(1 - p_D(\mathbf{x}_i))} \right) \left(\prod_{\mathbf{z} \in Z - W} \lambda c(\mathbf{z}) \right) \tag{3.50}$$

$$= e^{-\lambda} f_{\Upsilon(X)}(\emptyset) \left(\prod_{\mathbf{z} \in Z} \lambda c(\mathbf{z}) \right) \sum_{W \subseteq Z, |W| \leq n} \left(\sum_{\theta} \prod_{i:\theta(i)>0} \frac{p_D(\mathbf{x}_i).f_{k+1}(\mathbf{z}_{\theta(i)}|\mathbf{x}_i)}{(1 - p_D(\mathbf{x}_i))\lambda c(\mathbf{z}_{\theta(i)})} \right) \tag{3.51}$$

Therefore, the true multi-target likelihood probability function is rewritten as:

$$f_{k+1}(Z|X) = e^{\lambda} f_{k+1}(\emptyset|X) f_C(Z) \sum_{\theta} \left(\prod_{i:\theta(i)>0} \frac{p_D(\mathbf{x}_i).f_{k+1}(\mathbf{z}_{\theta(i)}|\mathbf{x}_i)}{(1 - p_D(\mathbf{x}_i))\lambda c(\mathbf{z}_{\theta(i)})} \right) \tag{3.52}$$

where the summation is taken over all the association functions (hypotheses) θ : $\{1, ..., n\} \rightarrow \{0, 1, ..., m\}$ and:

$$f_C(Z) = e^{-\lambda} \prod_{\mathbf{z} \in Z} \lambda c(\mathbf{z}) \tag{3.53}$$

$$f_{k+1}(\emptyset|X) = e^{-\lambda} \prod_{\mathbf{x} \in X} (1 - p_D(\mathbf{x})) \tag{3.54}$$

The next section derives the p.g.fl form of the likelihood function which is an important part in the derivation of PHD filter.

3.1.5 p.g.fls of Standard Measurement Model

Since the target generated measurement RFS ($\Upsilon(X)$) and the clutter RFS(C) are statistically independent RFS, from equation 2.95 the p.g.fl corresponding to the likelihood function can be given as:

$$G_{k+1}[h|X] = \int h^Z f_{k+1}(Z|X) \delta Z \tag{3.55}$$

$$= G_{\Upsilon(X)}[h|X].G_C[h] \tag{3.56}$$

The p.g.fl of the clutter RFS($G_C[h]$) can be found as follows:

$$G_C[h] = \int h^Z f_C(Z)\delta Z \tag{3.57}$$

$$= e^{-\lambda} \sum_{m=0}^{\infty} \frac{1}{m!} \int h(\mathbf{z}_1)...h(\mathbf{z}_m)\lambda^m c(\mathbf{z}_1)...c(\mathbf{z}_m)d\mathbf{z}_1...d\mathbf{z}_m \tag{3.58}$$

$$= e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m c[h]^m}{m!} \tag{3.59}$$

$$= e^{\lambda c[h]-\lambda} \tag{3.60}$$

where $c[h] = \int h(\mathbf{z})c(\mathbf{z})d\mathbf{z}$.
Since $\Upsilon(X) = \Upsilon(\mathbf{x}_1) \cup ... \cup \Upsilon(\mathbf{x}_n)$ where $\Upsilon(\mathbf{x}_1), ..., \Upsilon(\mathbf{x}_n)$ are statistically independent(assumption), from equation 2.95 it can be seen that:

$$G_{\Upsilon(X)}[h|X] = \prod_{i=1}^{n} G_{\Upsilon(\mathbf{x}_i)}[h] \tag{3.61}$$

$$= \prod_{i=1}^{n} (1 - p_D(\mathbf{x}_i) + p_D(\mathbf{x}_i) + p_D(\mathbf{x}_i).p_h(\mathbf{x}_i)) \tag{3.62}$$

where $p_h(\mathbf{x}_i) = \int h(\mathbf{z}) f_{k+1}(\mathbf{z}|\mathbf{x}_i)d\mathbf{z}$ Since $\Upsilon(\mathbf{x}_i)$ is a Bernoulli RFS, $G_{\Upsilon(\mathbf{x}_i)}[h]$ is obtained from equation 2.105.The p.g.fl of likelihood function can then be written as:

$$G_{k+1}[h|X] = (1 - p_D + p_D p_h)^X . e^{\lambda c[h]-\lambda} \tag{3.63}$$

where $(1 - p_D + p_D p_h)^X = \prod_{i=1}^{n} (1 - p_D(\mathbf{x}_i) + p_D(\mathbf{x}_i) + p_D(\mathbf{x}_i).p_h(\mathbf{x}_i))$.
The importance of p.g.fl can be seen here in the sense that p.g.fls greatly simplify the equations and give a compact form for likelihood function.These equations will be used in the derivation of the PHD filter (next chapter).

3.2 Multi-target Markov Density Function

The methodology to find the Markov density is very similar to the one followed in the previous section.This is because of the following similarities:

- Target disappearance is mathematically analogous to missed detection

- Target appearance is mathematically analogous to the generation of clutter measurements

Hence, the procedure to find the Markov density is exactly analogous to the one given in section 3.1.

- The first step is to define the RFS $\Xi_{k+1|k}(X) = \Gamma(X')$, describing the randomly varying state set at time t_{k+1} given that the target state set at t_k is X' and Γ is the RFS describing the multi-target state at t_k.

- The belief mass function (equation 2.38) corresponding to the RFS $\Xi_{k+1|k}$ completely characterizes statistics of the multi-target measurement set.

$$\beta_{k+1|k}(S|X') = Pr(\Xi_{k+1|k} \subseteq S|X') \tag{3.64}$$

- The probability density function of $\Xi_{k+1|k}$ i.e., $f_{k+1|k}(X|X')$ is then obtained from belief mass function(equation 2.75) as follows:

$$f_{k+1|k}(X|X') = \frac{\delta\beta_{k+1|k}}{\delta X}(\emptyset|X') \tag{3.65}$$

$$= \left[\frac{\delta\beta_{k+1|k}}{\delta X}(S|X')\right]_{S=\emptyset} \tag{3.66}$$

The probability density $f_{k+1|k}(X|X')$ is the *true multi-target Markov density function.*

3.2.1 Multi-target State Space

The multi-target state space is similar to the multi-target measurement space. *multi-target state* is best represented by a Random Finite Set since the states as well as the number of targets vary. Let X be a RFS representing the multi-target state, then:

$$X = \{\mathbf{x}_1, ..., \mathbf{x}_n\} \tag{3.67}$$

is a instantiation (a sample) of the RFS where $\mathbf{x}_1, ..., \mathbf{x}_n$ are elements of X which belong to a *single target state space* $\mathfrak{X}_0$. Then the various instantiations of X can be given by:

$$X = \emptyset \text{ (no targets present)} \tag{3.68}$$

$$X = \{\mathbf{x}_1\} \text{ (single target with state}\mathbf{x}_1 \text{ is present)} \tag{3.69}$$

$$X = \{\mathbf{x}_1, \mathbf{x}_2\} \text{ (two targets with states } \mathbf{x}_1 \neq \mathbf{x}_2 \text{ are present)} \tag{3.70}$$

and a sample with n targets is given by:

$$X = \{\mathbf{x}_1, ..., \mathbf{x}_n\} \text{ (n targets with states } \mathbf{x}_1 \neq ... \neq \mathbf{x}_n \text{ are present)} \tag{3.71}$$

Then the *multi-target state space* $\mathfrak{X}$ is the hyperspace of all the finite sets whose elements belong to the single-target state space $\mathfrak{X}_0$.

3.2.2 Standard Motion Model

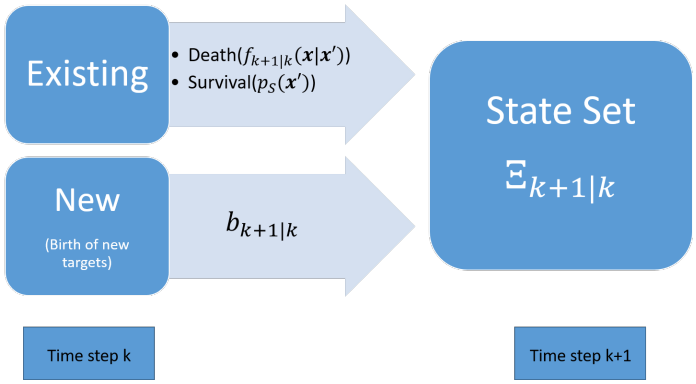


Figure 3.2. Standard State Model

The state set $\Xi_{k+1|k}$ can be divided into two parts where one part consists of states of the existing targets at time step t_k and the other part consists of the states of new objects entering the scene between t_k and t_{k+1}. Similar to the case of the measurement model, some assumptions to be made in order to derive the Markov density function $f_{k+1|k}(X|X')$. They are:

- The probability that a target with state $\mathbf{x}'$ at t_k will propagate to a state $\mathbf{x}$ at t_{k+1} is described by the single-target Markov transition density $f_{k+1|k}(\mathbf{x}|\mathbf{x}')$ which corresponds to the single-target motion model $\mathbf{X}_{k+1|k} = \Phi_k(\mathbf{x}') + \mathbf{V}_k(\mathbf{V}_k$ is process noise).

- A target with a state $\mathbf{x}'$ in the scene at t_k, is alive at the time step t_{k+1} with a probability $p_S(\mathbf{x}')$. Consequently, it is assumed that individual target state sets are Bernoulli RFS (section 2.8.1).

$$p_S(\mathbf{x}) \equiv p_S^{k+1|k}(\mathbf{x}') \tag{3.72}$$

- The appearance of new objects with state set X i.e., the birth objects is modeled as a RFS $B(X)$.

- It is also assumed that the target death, target birth and the target motion are statistically independent.

3.2.3 State RFS

Let $X' = \mathbf{x}'_1, ..., \mathbf{x}'_n$ be the multi-target state set at time step t_k. Then, the predicted state set(X) at t_{k+1} will be of the form:

$$\overset{\text{predicted state set}}{\Xi_{k+1|k}} = \overset{\text{existing targets set}}{\Gamma(X')} \cup \overset{\text{Birth Targets set}}{B} \tag{3.73}$$

i.e., if X is an instantiation of $\Xi_{k+1|k}$, then:

$$\overset{\text{predicted state set}}{X} = \overset{\text{existing targets set}}{\{\mathbf{x}_1, ..., \mathbf{x}_{n'}\}} \cup \overset{\text{Birth Targets set}}{\{\mathbf{b}_1, ..., \mathbf{b}_{m'}\}} \tag{3.74}$$

Furthermore, the existing targets set $\Gamma(X')$ can be expressed as a union of n RFS as follows:

$$\overset{\text{existing targets set}}{\Gamma(X')} = \overset{\text{state set from } \mathbf{x}'_1}{\Gamma(\mathbf{x}'_1)} \cup ... \cup \overset{\text{state set from } \mathbf{x}'_n}{\Gamma(\mathbf{x}'_n)} \tag{3.75}$$

As mentioned in the assumptions, $\Gamma(\mathbf{x}'_i), i = 1, ..., n$ are modeled as Bernoulli RFS and have the form:

$$\Gamma(\mathbf{x}'_i) = \emptyset^{p_S(\mathbf{x}'_i)} \cap \mathbf{X}_i \tag{3.76}$$

where $\mathbf{X}_i = \Phi(\mathbf{x}'_i), \mathbf{X} \in \mathfrak{X}_0$ is the single-target motion model and $\emptyset^{p_S(\mathbf{x}'_i)}$ is a discrete RFS(section 2.8.1) defined as follows.

$$Pr(\emptyset^{p_S(\mathbf{x}'_i)} = T) = \begin{cases} 1 - p_S(\mathbf{x}'_i) & \text{if } T = \emptyset \\ p_S(\mathbf{x}'_i) & \text{if } T = \mathfrak{X}_0 \\ 0 & \text{if otherwise} \end{cases} \tag{3.77}$$

The probability density function of $\Gamma(\mathbf{x}'_i)$ is then by equation 2.104 as:

$$f_{\Upsilon(\mathbf{x}'_i)}(X) = \begin{cases} p_S(\mathbf{x}'_i).f_{k+1|k}(\mathbf{x}|\mathbf{x}'_i) & \text{if } X = \{\mathbf{x}\} \\ 1 - p_S(\mathbf{x}'_i) & \text{if } X = \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{3.78}$$

The belief-mass function of $\Gamma(\mathbf{x}'_i)$ is then given by (equation 2.103):

$$\beta_{\Gamma(\mathbf{x}'_i)}(S) = 1 - p_S(\mathbf{x}'_i) + p_S(\mathbf{x}'_i) \int_S f_{k+1|k}(\mathbf{x}|\mathbf{x}'_i) d\mathbf{x} \tag{3.79}$$

Finally it is also assumed that $\Gamma(\mathbf{x}'_1), ..., \Gamma(\mathbf{x}'_n), B(X)$ are statistically independent RFS.

3.2.4 Derivation of Markov density function ($f_{k+1|k}(X|X')$)

If it is also assumed that the birth RFS $B(X)$ is assumed to be a Poisson RFS (Section 2.8.2, with a parameter μ_0 and elements distributed in space according to the density $b(\mathbf{x})$), then the motion model is exactly analogous to the measurement model. Then, if $B = \{\mathbf{b}_1, ..., \mathbf{b}_{m'}\}$, the cardinality distribution of B($|B| = m', m' \geq 0$) is:

$$p_{|B|}(m') = \frac{e^{-\mu}\mu^{m'}}{m'!}$$

and conditioned on $|B| = m'$, the elements $\mathbf{b}_1, ..., \mathbf{b}_{m'}$ are independent and identically distributed (i.i.d) random vectors with probability density $b(\mathbf{x}) = f_B(\mathbf{x})$.

Hence, the formal derivation of $f_{k+1|k}(X|X')$ is exactly the same as in Section 3.1.4. Therefore, the true Markov density can be written in analogous to equation 3.52.

$$f_{k+1|k}(X|X') = e^{\mu} f_{k+1|k}(\emptyset|X') f_B(X) \sum_{\theta} \left(\prod_{i:\theta(i)>0} \frac{p_S(\mathbf{x}'_i).f_{k+1|k}(\mathbf{x}_{\theta(i)}|\mathbf{x}'_i)}{(1-p_S(\mathbf{x}'_i))\mu_0 b(\mathbf{x}_{\theta(i)})} \right) \tag{3.80}$$

where the summation is taken over all the association functions(hypotheses) θ : $\{1,...,n'\} \rightarrow \{0,1,...,n\}$ and:

$$f_B(X) = e^{-\mu} \prod_{\mathbf{x}\in X} \mu b(\mathbf{x}) \tag{3.81}$$

$$f_{k+1|k}(\emptyset|X') = e^{-\mu} \prod_{\mathbf{x}'\in X'} (1-p_S(\mathbf{x}')) \tag{3.82}$$

Similarly, the p.g.fl of the Markov density can be given analogously to equation 3.63.

$$G_{k+1|k}[h|X'] = (1-p_S+p_S p_h)^{X'}.e^{\mu b[h]-\mu} \tag{3.83}$$

where $(1-p_S+p_S p_h)^{X'} = \prod_{i=1}^{n'}(1-p_S(\mathbf{x}'_i)+p_S(\mathbf{x}'_i)+p_S(\mathbf{x}'_i).p_h(\mathbf{x}'_i))$, $b[h] = \int h(\mathbf{x})b(\mathbf{x})d\mathbf{x}$, $p_h(\mathbf{x}'_i) = \int h(\mathbf{x}) f_{k+1|k}(\mathbf{x}|\mathbf{x}'_i)d\mathbf{x}$.

The above p.g.fl is true when the birth RFS is assumed to be a Poisson RFS. This assumption is not always required as will be seen in the case of PHD filter (Chapter 4). So, in the case when $B(X)$ is not a Poisson RFS, the p.g.fl is given as follows:

$$G_{k+1|k}[h|X'] = (1-p_S+p_S p_h)^{X'}.G_B[h] \tag{3.84}$$

where $G_B[h] = \int h^X b(X)\delta X$($b(X)$ is the probability density of $B(X)$) is the p.g.fl of the the birth RFS. This equation is a direct implementation of equation 2.95 and using the fact that $\Gamma(\mathbf{x}'_i)$ is a Bernoulli RFS.

3.3 Multi-target Bayes Filter

The previous sections gave a formal derivation of the true Markov and true likelihood density. The multi-target Bayes filter can then be defined analogously to the single target case(Section 2.1).Let $Z^(k) : Z_1,...,Z_k$ be the measurement sets obtained

till time t_k. If $f_{k+1|k}(X|X')$(equation 3.80) and $f_{k+1}(Z|X))$(equation 3.52) are the multi-target Markov Density and the likelihood function respectively, then the predictor equation for multi-target Bayes filter is the analog of equation 2.5 and is given by:

$$f_{k+1|k}(X|Z^{(k)}) = \int f_{k+1|k}(X|X').f_{k|k}(X'|Z^{(k)})\delta X' \tag{3.85}$$

and the corrector equation is an analog to equation 2.6 and is given by:

$$f_{k+1|k+1}(X|Z^{(k+1)}) = \frac{f_{k+1}(Z_{k+1}|X).f_{k+1|k}(X|Z^{(k)})}{f_{k+1}(Z^{k+1}|Z^{(k)})} \tag{3.86}$$

where the *Bayes normalization factor* is given by:

$$f_{k+1}(Z^{k+1}|Z^{(k)}) = \int f_{k+1}(Z_{k+1}|X).f_{k+1|k}(X|Z^{(k)})\delta X \tag{3.87}$$

This filter then propagates through time as:

$$f_{0|0}(X|Z^{(0)}) \rightarrow f_{1|0}(X|Z^{(0)}) \rightarrow f_{1|1}(X|Z^{(1)}) \rightarrow$$
$$... \rightarrow f_{k|k}(X|Z^{(k)}) \rightarrow f_{k+1|k}(X|Z^{(k)}) \rightarrow f_{k+1|k+1}(X|Z^{(k+1)}) \rightarrow ...$$

3.4 p.g.fl Multi-target Bayes Filter

The p.g.fl form of the Bayes filter is introduced here since it is fundamental for the derivation of the PHD filter which is the topic of this thesis and will be discussed in the next chapter.

3.4.1 p.g.fl Multi-target Predictor

The p.g.fl corresponding to the predictor posterior $f_{k+1|k}(X|Z^{(k)})$ is defined as:

$$G_{k+1|k}[h] = \int h^X.f_{k+1|k}(X|Z^{(k)})\delta X \equiv \int h^X.f_{k+1|k}(X)\delta X \tag{3.88}$$

where $f_{k+1|k}(X|Z^{(k)})$ is obtained from equation 3.85.Then:

$$G_{k+1|k} = \int h^X . f_{k+1|k}(X)\delta X \tag{3.89}$$

$$= \int h^X . \Big(\int f_{k+1|k}(X|X') . f_{k|k}(X')\delta X' \Big) \delta X \tag{3.90}$$

$$= \int \Big(\int h^X . f_{k+1|k}(X|X')\delta X \Big) . f_{k|k}(X')\delta X' \tag{3.91}$$

$$= \int G_{k+1|k}[h|X'] . f_{k|k}(X')\delta X' \tag{3.92}$$

Using equation 3.84,

$$G_{k+1|k}[h] = G_B[h] \int (1 - p_S + p_S p_h)^{X'} . f_{k|k}(X')\delta X' \tag{3.93}$$

$$= G_B[h] . G_{k|k}[1 - p_S + p_S p_h] \tag{3.94}$$

The predicted multi-target posterior density can be obtained from it's p.g.fl as (given by equation 2.78) as:

$$f_{k+1|k}(X|Z^{(k)}) = \frac{\delta G_{k+1|k}}{\delta X}[0] \tag{3.95}$$

3.4.2 p.g.fl Multi-target Corrector

Consider the functional $F[g, h]$ defined as follows:

$$F[g, h] = \int h^X . G_{k+1}[g|X] . f_{k+1|k}(X)\delta X \tag{3.96}$$

where $G_{k+1}[g|X]$ is the p.g.fl of the likelihood function, which is:

$$G_{k+1}[g|X] = \int g^Z . f_{k+1}(Z|X)\delta Z \tag{3.97}$$

From equation 3.63,it is seen that:

$$G_{k+1}[g|X] = (1 - p_D + p_D p_g)^X . e^{\lambda c[g] - \lambda} \tag{3.98}$$

where $c[g] = \int g(\mathbf{z})c(\mathbf{z})d\mathbf{z}, p_g(x) = \int g(\mathbf{z}).f_{k+1}(\mathbf{z}|x)d\mathbf{z}$.Then $F[g,h]$ can be given as follows:

$$F[g,h] = \int h^X.G_{k+1}[g|X].f_{k+1|k}(X)\delta X \tag{3.99}$$

$$= e^{\lambda c[g]-\lambda} \int h^X.(1-p_D+p_Dp_g)^X.f_{k+1|k}(X)\delta X \tag{3.100}$$

$$= e^{\lambda c[g]-\lambda} \int [h.(1-p_D+p_Dp_g)]^X.f_{k+1|k}(X)\delta X \tag{3.101}$$

$$= e^{\lambda c[g]-\lambda}.G_{k+1|k}[h.(q_D+p_Dp_g)] \tag{3.102}$$

where $q_D(\mathbf{x}) = 1 - p_d(\mathbf{x})$.

Using the Radon-Nikod'ym theorem for functional derivatives (equation 2.77), the derivative $\frac{\delta G_{k+1}}{\delta Z}$ can be written as:

$$\frac{\delta G_{k+1}}{\delta Z}[g|X] = \int g^W.f_{k+1}(Z \cup W|X)\delta W \tag{3.103}$$

Then,

$$\frac{\delta F}{\delta Z}[g,h] = \int h^X.\left(\int g^W.f_{k+1}(Z \cup W|X)\delta W \right).f_{k+1|k}(X)\delta X \tag{3.104}$$

$$= \int g^W.\left(\int h^X.f_{k+1}(Z \cup W|X).f_{k+1|k}(X)\delta X \right)\delta W \tag{3.105}$$

This implies,

$$\frac{\delta F}{\delta Z}[0,h] = \int h^X.f_{k+1}(Z|X).f_{k+1|k}(X)\delta X \tag{3.106}$$

$$\frac{\delta F}{\delta Z}[0,1] = \int f_{k+1}(Z|X).f_{k+1|k}(X)\delta X = f_{k+1}(Z) \tag{3.107}$$

The p.g.fl corresponding to the corrected posterior probability distribution ($f_{k+1|k+1}(X|Z^{(k+1)})$, defined as in equation 3.86) can be written as:

$$G_{k+1|k+1}[h] = \int h^X.f_{k+1|k+1}(X)\delta X \tag{3.108}$$

$$= \int h^X.\frac{f_{k+1}(Z_{k+1}|X).f_{k+1|k}(X|Z^{(k)})}{f_{k+1}(Z^{k+1}|Z^{(k)})}\delta X \tag{3.109}$$

$$= \int h^X.\frac{f_{k+1}(Z|X).f_{k+1|k}(X))\delta X}{f_{k+1}(Z)} \tag{3.110}$$

$$= \frac{\frac{\delta F}{\delta Z_{k+1}}[0,h]}{\frac{\delta F}{\delta Z_{k+1}}[0,1]} \tag{3.111}$$

Equations 3.94 and 3.111 form the p.g.fl form of the Multi-target Bayes filter.

The Multi-target Bayes filter equations (equations 3.85 and 3.86), involve multiple complex integrals since they are composed of set integrals and this makes the filter equations computationally intractable especially when the number of targets are high. Also, the evaluation of the multi-target Markov density (equation 3.80) and the likelihood function (equation 3.52) involve combinatorial calculations which are computationally very expensive. The Bayes filter in its original form can still be applied by a Monte-Carlo approximation technique for special cases where the target motion is not complicated and the number of target are less (refer [37]).

The PHD filter is a first-order approximation of the Bayes filter, was developed to overcome the computational intractability of the multi-target Bayes filter. The PHD filter propagates only the first order statistical moment of the posterior density known as the *Probability Hypothesis Density* (PHD) instead of propagating the complete posterior density. This approximation is analogous to the Kalman Filter (section 2.2) where the first and the second moment of the density were propagated instead of the complete probability distribution. The next section formally derives the PHD filter.

4. PHD FILTER

4.1 Probability Hypothesis Density (PHD)

Intuitively, PHD can be thought of as the multi-target counterpart of an expected value.But the formally,PHD of a RFS Ξ is given by:

$$D_{\Xi}(\mathbf{x}) = E[\delta_{\Xi}(x)] = \int \delta_X(x) f_{\Xi}(X) \delta X \tag{4.1}$$

where $x \in \mathfrak{X}_0$,the single target state space and where:

$$\delta_X(\mathbf{x}) = \begin{cases} 0 & \text{if } X = \emptyset \\ \sum_{w \in X} \delta_w(\mathbf{x}) & \text{otherwise} \end{cases} \tag{4.2}$$

It must be noted that PHD is not a probability density but rather a density function.The integral of $D_{\Xi}(\mathbf{x})$ in any region S will result in the expected number of objects with state $\mathbf{x}$ in that region i.e.,

$$\int_S D_{\Xi}(\mathbf{x}) d\mathbf{x} = E[\ |S \cap \Xi|\ ] \tag{4.3}$$

where $\mid S \cap \Xi \mid$ gives the number of objects or number of elements of Ξ in region S.Equation 4.3 is obtained as follows:

$$E[\ |S \cap \Xi|\ ] = \int |X \cap S| f_{\Xi}(X) \delta X \tag{4.4}$$

$$= \int \left(\sum_{\mathbf{x} \in X} \mathbb{1}_S(\mathbf{x}) \right) f_{\Xi}(X) \delta X \tag{4.5}$$

$$= \sum_{n \geq 0} \frac{1}{n!} \underbrace{\int ... \int}_{n} (\mathbb{1}_S(\mathbf{x}_1) + ... + \mathbb{1}_S(\mathbf{x}_n)).f_{\Xi}(X) d\mathbf{x}_1 ... d\mathbf{x}_n \tag{4.6}$$

$$= \sum_{n \geq 0} \frac{n}{n!} \underbrace{\int ... \int}_{n} \mathbb{1}_S(\mathbf{x}_n).f_{\Xi}(\{\mathbf{x}_1, ..., \mathbf{x}_n\}) d\mathbf{x}_1 ... d\mathbf{x}_n \tag{4.7}$$

$$= \int_S \left(\sum_{i \geq 0} \frac{1}{i!} \underbrace{\int ... \int}_{i-1} f_{\Xi}(\{\mathbf{x}_1, ..., \mathbf{x}_{i-1}, \mathbf{x}_i\}) d\mathbf{x}_1 ... d\mathbf{x}_{i-1} \right) d\mathbf{x}_i \tag{4.8}$$

$$= \int_S D_{\Xi}(\mathbf{x}_i) d\mathbf{x}_i = \int_S D_{\Xi}(\mathbf{x}) d\mathbf{x} \tag{4.9}$$

The last equation is obtained from the relationship between the PHD and probability density function which is shown in the following section.Hence, integrating D_{Ξ} over the entire space i.e., $\int_{\mathbb{x}_0} D_{\Xi}(\mathbf{x}) d\mathbf{x}$ will not give 1 (property of a probability density function) always but rather give the expected number of objects with state/element $\mathbf{x}$ per unit volume.

PHD is also known as the intensity density or the first-moment density.The following sections show the relationships between the PHD,probability distribution function and p.g.fl.

4.1.1 Formulae for PHD

An alternate definition of PHD of a RFS Ξ can be derived in the following manner(refer [12],chapter 16):

$$D_{\Xi}(\mathbf{x}) = \int \delta_X(x) f_{\Xi}(X) \delta X \tag{4.10}$$

$$= \sum_{n \geq 1} \frac{1}{n!} \underbrace{\int ... \int}_{n} [\delta_{\mathbf{x}_1}(\mathbf{x}) + ... + \delta_{\mathbf{x}_n}(\mathbf{x})] f_{\Xi}(\mathbf{x}_1, ..., \mathbf{x}_n) d\mathbf{x}_1 ... d\mathbf{x}_n \tag{4.11}$$

$$= \sum_{n \geq 1} \frac{n}{n!} \underbrace{\int ... \int}_{n-1} f_{\Xi}(x, \mathbf{w}_1, ..., \mathbf{w}_{n-1}) d\mathbf{w}_1 ... d\mathbf{w}_{n-1} \tag{4.12}$$

The last equation is obtained from:

$$\underbrace{\int ... \int}_{n} \delta_{\mathbf{x}_i} f_{\Xi}(\mathbf{x}_1, ..., \mathbf{x}_n) d\mathbf{x}_1 ... d\mathbf{x}_n = \underbrace{\int ... \int}_{n-1} f_{\Xi}(\mathbf{x}_1, ..., \mathbf{x}..., \mathbf{x}_{n-1}) d\mathbf{x}_1 ... d\mathbf{x}_{n-1} \tag{4.13}$$

Then,

$$D_{\Xi}(\mathbf{x}) = \sum_{n \geq 1} \frac{1}{(n-1)!} \underbrace{\int ... \int}_{n-1} f_{\Xi}(x, \mathbf{w}_1, ..., \mathbf{w}_{n-1}) d\mathbf{w}_1 ... d\mathbf{w}_{n-1} \tag{4.14}$$

$$= \sum_{i \geq 0} \frac{1}{i!} \underbrace{\int ... \int}_{i} f_{\Xi}(\mathbf{x}, \mathbf{w}_1, ..., \mathbf{w}_i) d\mathbf{w}_1 ... d\mathbf{w}_i \tag{4.15}$$

$$= \int f_{\Xi}(\{\mathbf{x}\} \cup W) \delta W \tag{4.16}$$

From the Radon-Nikodým theorem for functional differentiation it can be seen that(equation 2.78):

$$\frac{\delta G_{\Xi}}{\delta X}[h] = \int h^W . f_{\Xi}(X \cup W) \delta W \tag{4.17}$$

Substituting $h = 1, X = \{\mathbf{x}\}$,

$$\frac{\delta G_{\Xi}}{\delta \mathbf{x}}[1] = \int f_{\Xi}(\{\mathbf{x}\} \cup W) \delta W = D_{\Xi}(\mathbf{x}) \tag{4.18}$$

4.1.2 PHD of a Poisson RFS

As was seen in the previous chapter, the modeling of the standard measurement and the standard motion model included the application of Poisson RFS.This section

formally derives the PHD of a Poisson RFS.

The PHD of a Poisson RFS(section 2.8.2) Ξ with parameter λ and distribution $f(\mathbf{x})$ is obtained by applying the equation 4.18 on the p.g.fl of a Poisson RFS(equation 2.121).The first functional derivative of the p.g.fl is given as(using the chain rule(equation 2.83) and the derivative of a linear functional(equation 2.63)):

$$\frac{\delta G_{\Xi}}{\delta \mathbf{x}}[h] = e^{\lambda f[h]-\lambda}.\frac{\delta}{\delta \mathbf{x}}\lambda f[h] = e^{\lambda f[h]-\lambda}.\lambda.f(\mathbf{x}) \tag{4.19}$$

where $f[h] = \int h(\mathbf{x})f(\mathbf{x})d\mathbf{x}$.Then it's PHD is given by:

$$D_{\Xi}(\mathbf{x}) = \frac{\delta G_{\Xi}}{\delta \mathbf{x}}[1] = \lambda.f(\mathbf{x}) \tag{4.20}$$

Note that $\frac{\delta G_{\Xi}}{\delta \mathbf{x}}[1] \equiv \frac{\delta G_{\Xi}[h]}{\delta \mathbf{x}}\Big|_{h=1}$

4.2 The PHD Filter

The following sections detail the assumptions underlying the predictor and corrector parts of the PHD filter and then formally derive the filter equations in each section. The derivation presented in the following sections was obtained from [12].

4.2.1 PHD Filter Predictor

Assumptions

The PHD filter assumes all the assumptions that were described in the standard multi-target motion model (section 3.2.2) i.e.,the targets are assumed to move independent of each other,existing targets can disappear form the scene(p_S) and new targets can enter the scene(birth) and these processes are independent of each other.The overall assumptions can be stated as follows:

- $f_{k+1|k}(\mathbf{x}|\mathbf{x}')$ is the single target Markov transition density.
- $p_S(\mathbf{x}') \equiv p_{S,k+1|k}(\mathbf{x}')$ is the probability that a target with state $\mathbf{x}'$ at time t_k will survive at time t_{k+1}

- $B_{k+1|k}(X)$ is the RFS that describes the new objects entering the scene between time t_{k+1} and t_k with the state set X.It's PHD is then given by:

$$b_{k+1|k}(\mathbf{x}) = \int b_{k+1|k}(\{x\} \cup W)\delta W = \frac{\delta G_B}{\delta \mathbf{x}}[1] \tag{4.21}$$

  where $b_{k+1|k}(X)$ is the multi-object probability density of the birth RFS and $G_B[h]$ is the p.g.fl of the birth RFS.

The following derivations for the predictor and corrector equations can also be found in [38].

Predictor Equations

Based on the above assumptions, the p.g.fl of the prediction posterior density can be written as(equation 3.94):

$$G_{k+1|k}[h|X'] = G_B[h].G_{k|k}(1 - p_S + p_S p_h) \tag{4.22}$$

Let $T[h(x)] = 1-p_S+p_S p_h$ where $1-p_S+p_S p_h = 1-p_S(\mathbf{x})+p_S(\mathbf{x}) \int h(\mathbf{x}')f_{k+1|k}(\mathbf{x}'|\mathbf{x})d\mathbf{x}'$. Then,

$$G_{k+1|k}[h|X'] = G_B[h].G_{k|k}(T[h]) \tag{4.23}$$

The PHD filter prediction equation is then obtained by using the relationship in equation 4.18 i.e.,

$$D_{k|k+1}(\mathbf{x}) = \frac{\delta G_{k|k+1}}{\delta \mathbf{x}}[1] \tag{4.24}$$

Using the product rule(equation 2.82),

$$D_{k+1|k}(\mathbf{x}) = \frac{\delta G_B}{\delta \mathbf{x}}[1].G_{k|K}[T[1]] + G_B[1].\frac{\delta G_{k|k}}{\delta \mathbf{x}}[T[1]] \tag{4.25}$$

In order to evaluate the above expression the following chain rule(detailed proof can be found in [12],page 806,807) will be used:

$$\frac{\delta}{\delta \mathbf{x}}G_{k|K}[T[h]] = \int \frac{\delta T}{\delta \mathbf{x}}[h](\mathbf{w}).\frac{\delta G_{k|k}}{\delta \mathbf{w}}[T[h]]d\mathbf{w} \tag{4.26}$$

Then,

$$\frac{\delta T}{\delta \mathbf{x}}[h](w) = p_S(\mathbf{w}).f_{k+1|k}(\mathbf{x}|\mathbf{w}) \tag{4.27}$$

The above equation is obtained by the application of the linear rule and the constant rule(equations 2.79 and 2.80).Also, $T[1] = 1 - p_S + p_S \int f_{k+1|k}(\mathbf{x}'|\mathbf{x})d\mathbf{x}' = 1$.Then,

$$\frac{\delta}{\delta \mathbf{x}}G_{k|K}[T[1]] = \int p_S(\mathbf{w}).f_{k+1|k}(\mathbf{x}|\mathbf{w}).\frac{\delta G_{k|k}}{\delta \mathbf{w}}[1]d\mathbf{w} \tag{4.28}$$

$$= \int p_S(\mathbf{w}).f_{k+1|k}(\mathbf{x}|\mathbf{w}).D_{k|k}(\mathbf{w})d\mathbf{w} \tag{4.29}$$

Note that for any p.g.fl G,$G[1] = \int f\Psi(Y)\delta Y = 1$.Therefore, substituting all the quantities in equation 4.25:

$$D_{k+1|k}(\mathbf{x}) = b_{k+1|k}(\mathbf{x}).1 + 1. \int p_S(\mathbf{w})f_{k+1|k}(\mathbf{x}|\mathbf{w})D_{k|k}(\mathbf{w})d\mathbf{w} \tag{4.30}$$

$$= b_{k+1|k}(\mathbf{x}) + \int p_S(\mathbf{x}')f_{k+1|k}(\mathbf{x}|\mathbf{x}')D_{k|k}(\mathbf{x}')d\mathbf{x}' \tag{4.31}$$

Equation 4.31 is the PHD filter predictor equation.

4.2.2 PHD Filter Corrector

Similar to the prediction step, PHD filter assumes all the assumptions that were made in the standard measurement model(section 3.1.2) i.e., no target generates more than a single measurement, target generated measurement process is statistically independent of clutter measurement process etc.They can be stated more precisely as follows:

- The likelihood function is given by $L_{\mathbf{z}}(\mathbf{x}) = f_{k+1}(\mathbf{z}|\mathbf{x})$ i.e., it is the probability that a target with state $\mathbf{x}$ generates a measurement $\mathbf{z}$.
- $p_D(\mathbf{x}) \equiv p_{D,k+1}(x)$ is the probability that a target with state $\mathbf{x}$ is detected at time step t_{k+1}.
- $C(Z)$ is the RFS that describes the clutter measurement set Z.It is also assumed that C is a Poisson RFS(section 2.8.2) with parameter $\lambda \equiv \lambda_{k+1}$ and the spatial distribution of the elements $c(\mathbf{z}) \equiv c_{k+1}(\mathbf{z})$.

- An additional assumption made in the PHD filter is made on the predicted posterior density i.e., it is assumed that $f_{k+1|k}(X|Z)$ is Poisson.More precisely, the predicted multi-target RFS governed by $f_{k+1|k}(X|Z)$ is assumed to be Poisson RFS.This is a reasonable approximation when the interactions between objects are negligible(refer [38]).Then the PHD of $f_{k+1|k}(X|Z^{(k)})$ is $D_{k|k+1}(\mathbf{x}) = \mu s(\mathbf{x})$.And also,

$$G_{k+1|k}[h] = e^{\mu s[h]-\mu}, \quad s[h] = \int h(\mathbf{x})s(\mathbf{x})d\mathbf{x} \tag{4.32}$$

  where $\mu, s(\mathbf{x})$ are the Poisson parameters.

Corrector Equations

Based on the assumptions above, the p.g.fl of $f_{k+1|k+1}(X|Z^{(k+1)})$ is given by equation 3.111, which is:

$$G_{k+1|k+1}[h] = \frac{\frac{\delta F}{\delta Z_{k+1}}[0,h]}{\frac{\delta F}{\delta Z_{k+1}}[0,1]} \tag{4.33}$$

Then using the relationship between PHD and p.g.fl(equation 4.18),PHD of $f_{k+1|k+1}(X|Z^{(k+1)})$ is given by:

$$D_{k+1|k+1}(\mathbf{x}) = \frac{\frac{\delta^{m+1}F}{\delta \mathbf{z}_m ... \delta \mathbf{z}_1 \delta \mathbf{x}}[0,h]\Big|_{h=1}}{\frac{\delta^m F}{\delta \mathbf{z}_m ... \delta \mathbf{z}_1}[0,1]} \tag{4.34}$$

where $Z^{(k+1)} = \{\mathbf{z}_1, ..., \mathbf{z}_m\}$.From equation 3.102,it can be seen that:

$$F[g,h] = e^{\lambda c[g]-\lambda}.G_{k+1|k}[h.(q_D + p_D p_g)] \tag{4.35}$$

Then,

$$F[g,h] = \exp\left(\lambda c[g] - \lambda + \mu s[hq_D] + \mu s[hp_D p_g] - \mu\right) \tag{4.36}$$

$$F[g,1] = \exp\left(\lambda c[g] - \lambda - \mu s[p_D] + \mu s[p_D p_g]\right) \tag{4.37}$$

The denominator in equation 4.34 for $Z^{(k+1)} = \emptyset$ is then given by:

$$F[0,1] = e^{-\lambda - \mu s[p_D]} \tag{4.38}$$

For $Z^{(k+1)} = \{\mathbf{z}_1\}$,using the functional derivative rules(product rule(equation 2.82),linear rule (equation 2.80)):

$$\frac{\delta F}{\delta \mathbf{z}_1}[g,1] = F[g,1].\frac{\delta}{\delta \mathbf{z}_1}(\lambda c[g] - \lambda - \mu s[p_D] + \mu s[p_D p_g])) \tag{4.39}$$

$$= F[g,1].(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}])) \tag{4.40}$$

$$\frac{\delta F}{\delta \mathbf{z}_1}[0,1] = e^{-\lambda - \mu s[p_D]}.(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}])) \tag{4.41}$$

where $L_{\mathbf{z}_1} = f_{k+1}(\mathbf{z}_1|\mathbf{x})$. Then if $Z^{(k+1)} = \{\mathbf{z}_1, \mathbf{z}_2\}$,

$$\frac{\delta^2 F}{\delta \mathbf{z}_2 \delta \mathbf{z}_1}[g,1] = \left(\frac{\delta}{\delta \mathbf{z}_2} F[g,1]\right).(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}])) \tag{4.42}$$

$$= F[g,1].(\lambda c(\mathbf{z}_2) + \mu s[p_D L_{\mathbf{z}_2}])).(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}])) \tag{4.43}$$

$$\frac{\delta^2 F}{\delta \mathbf{z}_2 \delta \mathbf{z}_1}[0,1] = e^{-\lambda - \mu s[p_D]}.(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}])).(\lambda c(\mathbf{z}_2) + \mu s[p_D L_{\mathbf{z}_2}])) \tag{4.44}$$

Using a similar process for $Z^{(k+1)} = \{\mathbf{z}_1 ..., \mathbf{z}_m\}$,

$$\frac{\delta^m F}{\delta \mathbf{z}_m ... \delta \mathbf{z}_1}[0,1] = e^{-\lambda - \mu s[p_D]}.(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}]))...(\lambda c(\mathbf{z}_m) + \mu s[p_D L_{\mathbf{z}_m}])) \tag{4.45}$$

The numerator in equation 4.34 is obtained by first taking the functional derivative of $F[g,h]$ w.r.t $\mathbf{x}$.Then,

$$\frac{\delta F}{\delta \mathbf{x}}[g,h] = F[g,h].\frac{\delta}{\delta \mathbf{x}}(\lambda c[g] - \lambda + \mu s[h q_D] + \mu s[h p_d p_g] - \mu) \tag{4.46}$$

$$= F[g,h].(\mu q_D(\mathbf{x}) s(\mathbf{x}) + \mu p_D(\mathbf{x}) p_g(\mathbf{x}) s(\mathbf{x})) \tag{4.47}$$

$$\frac{\delta F}{\delta \mathbf{x}}[g,1] = F[g,1].(\mu q_D(\mathbf{x}) s(\mathbf{x}) + \mu p_D(\mathbf{x}) p_g(\mathbf{x}) s(\mathbf{x})) \tag{4.48}$$

The derivative of the numerator is evaluated in a similar manner to the one used to evalute the derivative of the denominator.Consider $Z = \emptyset$,then:

$$\frac{\delta^{m+1} F}{\delta \mathbf{z}_m ... \delta \mathbf{z}_1 \delta \mathbf{x}}[0,h]\bigg|_{h=1} = \frac{\delta F}{\delta \mathbf{x}}[0,1] = F[0,1].\mu q_D(\mathbf{x}) s(\mathbf{x}) \tag{4.49}$$

Similarly, for $Z = \{\mathbf{z}_1\}$:

$$\frac{\delta^2 F}{\delta \mathbf{z}_1 \delta \mathbf{x}}[g,1] = F[g,1].(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}]).(\mu q_D(\mathbf{x}) s(\mathbf{x}) + \mu p_D(\mathbf{x}) p_g(\mathbf{x}) s(\mathbf{x})) \tag{4.50}$$

$$+ F[g,1].\mu p_D(\mathbf{x}) L_{\mathbf{z}_1}(\mathbf{x}) s(\mathbf{x})$$

Then,

$$\frac{\delta^{m+1}F}{\delta\mathbf{z}_m...\delta\mathbf{z}_1\delta\mathbf{x}}[0,h]\bigg|_{h=1} = \frac{\delta^2 F}{\delta\mathbf{z}_1\delta\mathbf{x}}[0,1] \tag{4.51}$$

$$= e^{-\lambda-\mu s[p_D]}.(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}]).\mu q_D(\mathbf{x})s(\mathbf{x}) \tag{4.52}$$

$$+ e^{-\lambda-\mu s[p_D]}.\mu p_D(\mathbf{x})L_{\mathbf{z}_1}(\mathbf{x})s(\mathbf{x})$$

If $Z = \{\mathbf{z}_1, \mathbf{z}_2\}$,

$$\frac{\delta^3 F}{\delta\mathbf{z}_2\delta\mathbf{z}_1\delta\mathbf{x}}[g,1] = \left(\frac{\delta}{\delta\mathbf{z}_2}F[g,1]\right).(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}]) \tag{4.53}$$

$$.(\mu q_D(\mathbf{x})s(\mathbf{x}) + \mu p_D(\mathbf{x})p_g(\mathbf{x})s(\mathbf{x}))$$

$$+ F[g,1].(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}])$$

$$.\frac{\delta}{\delta\mathbf{z}_2}(\mu q_D(\mathbf{x})s(\mathbf{x}) + \mu p_D(\mathbf{x})p_g(\mathbf{x})s(\mathbf{x}))$$

$$+ \left(\frac{\delta}{\delta\mathbf{z}_2}F[g,1]\right).\mu p_D(\mathbf{x})L_{\mathbf{z}_1}(\mathbf{x})s(\mathbf{x})$$

$$= F[g,1].(\lambda c(\mathbf{z}_2) + \mu s[p_D L_{\mathbf{z}_2}]) \tag{4.54}$$

$$.(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}])$$

$$.(\mu q_D(\mathbf{x})s(\mathbf{x}) + \mu p_D(\mathbf{x})p_g(\mathbf{x})s(\mathbf{x}))$$

$$+ F[g,1].(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}])$$

$$.\mu p_D(\mathbf{x})L_{\mathbf{z}_2}(\mathbf{x})s(\mathbf{x})$$

$$+ F[g,1].(\lambda c(\mathbf{z}_2) + \mu s[p_D L_{\mathbf{z}_2}]).$$

$$.\mu p_D(\mathbf{x})L_{\mathbf{z}_1}(\mathbf{x})s(\mathbf{x})$$

Therefore,for $Z = \{\mathbf{z}_1, \mathbf{z}_2\}$,

$$\frac{\delta^{m+1}F}{\delta\mathbf{z}_m ... \delta\mathbf{z}_1 \delta\mathbf{x}}[0,h]\bigg|_{h=1} = \frac{\delta^3 F}{\delta\mathbf{z}_2\delta\mathbf{z}_1\delta\mathbf{x}}[0,1] \tag{4.55}$$

$$\begin{aligned} &= e^{-\lambda-\mu s[p_D]}.(\lambda c(\mathbf{z}_2) + \mu s[p_D L_{\mathbf{z}_2}]) \\ &.(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}]).\mu q_D(\mathbf{x}) s(\mathbf{x}) \\ &+ e^{-\lambda-\mu s[p_D]}.(\lambda c(\mathbf{z}_1) + \mu s[p_D L_{\mathbf{z}_1}]) \\ &.\mu p_D(\mathbf{x}) L_{\mathbf{z}_2}(\mathbf{x}) s(\mathbf{x}) \\ &+ e^{-\lambda-\mu s[p_D]}.(\lambda c(\mathbf{z}_2) + \mu s[p_D L_{\mathbf{z}_2}]). \\ &.\mu p_D(\mathbf{x}) L_{\mathbf{z}_1}(\mathbf{x}) s(\mathbf{x}) \end{aligned} \tag{4.56}$$

Hence, by using a proof of induction,for $Z_{k+1} = \{\mathbf{z}_1, ..., \mathbf{z}_m\}$ it can shown that,

$$\begin{aligned} \frac{\delta^{m+1}F}{\delta\mathbf{z}_m ... \delta\mathbf{z}_1 \delta\mathbf{x}}[0,h]\bigg|_{h=1} &= e^{-\lambda-\mu s[p_D]}.\prod_{\mathbf{z}\in Z_{k+1}}(\lambda c(\mathbf{z}) + \mu s[p_D L_{\mathbf{z}}]).\mu q_D(\mathbf{x}) s(\mathbf{x}) \\ &+ e^{-\lambda-\mu s[p_D]}.\left(\prod_{\mathbf{z}\in Z_{k+1}}(\lambda c(\mathbf{z}) + \mu s[p_D L_{\mathbf{z}}])\right) \\ &+ \sum_{\mathbf{z}\in Z_{k+1}} \frac{\mu p_D(\mathbf{x}) L_{\mathbf{z}}(\mathbf{x}) s(\mathbf{x})}{\lambda c(\mathbf{z}) + \mu s[p_D L_{\mathbf{z}}]} \end{aligned} \tag{4.57}$$

Substituting equations 4.45,4.57 in equation 4.34,it can be seen that:

$$D_{k+1|k+1}(\mathbf{x}) = \left(1 - p_D(\mathbf{x}) + \sum_{\mathbf{z}\in Z_{k+1}} \frac{p_D(\mathbf{x}) L_{\mathbf{z}}(\mathbf{x})}{\lambda c(\mathbf{z}) + \mu D_{k+1|k}[p_D L_{\mathbf{z}}]}\right) D_{k+1|k}(\mathbf{x}) \tag{4.58}$$

Note that $D_{k+1|k}(\mathbf{x}) = \mu s(\mathbf{x})$,this was obtained as a result of the assumption that the RFS described by $f_{k+1|k}(X)$ is a Poisson RFS.Hence ,given an initial PHD $D_{0|0}(\mathbf{x})$, associated with the initial multi-object probability density $f_{0|0}(X)$,the posterior PHD can be propagated in time via the following PHD recursion equations:

$$D_{k+1|k}(\mathbf{x}) = b_{k+1|k}(\mathbf{x}) + \int p_S(\mathbf{x}') f_{k+1|k}(\mathbf{x}|\mathbf{x}') D_{k|k}(\mathbf{x}') d\mathbf{x}' \tag{4.59}$$

$$D_{k+1|k+1}(\mathbf{x}) = \left(1 - p_D(\mathbf{x}) + \sum_{\mathbf{z}\in Z_{k+1}} \frac{p_D(\mathbf{x}) L_{\mathbf{z}}(\mathbf{x})}{\lambda c(\mathbf{z}) + \mu D_{k+1|k}[p_D L_{\mathbf{z}}]}\right) D_{k+1|k}(\mathbf{x}) \tag{4.60}$$

Since the PHD is a first order statistical moment of a RFS, there is a loss of information if the PHD is propagated instead of the complete multi-object probability distribution.However, comparing the multi-target Bayes filter(equations 3.85 and 3.86)

and the PHD recursive filter(equations 4.31 and 4.58),it is clear that the PHD filter completely circumvents the combinatorial computations arising in the multi-target Bayes filter.Moreover, since the PHD filter operates on the single target state space $\mathfrak{X}_0$ it is computationally less expensive than the Bayes filter which operates on sets. Even though the complexity is greatly reduced by the PHD filter, there are no closed form solutions available in general and the numerical integrations suffer from the "curse of dimensionality".Nonetheless, under certain additional assumptions the PHD recursion (equations 4.31 and 4.58) admits closed form solution.The next section deals with the derivation of the *Gaussian Mixture Probability Hypothesis Density*(GM-PHD) filter which is the topic of this thesis.

4.3 GM-PHD Filter

It will be seen that for a certain class of multi-target models,the PHD recursion admits a closed form solution. As discussed previously, the GM-PHD filter was developed by [1]. The following section details the additional assumptions (refer [1]) that are necessary to obtain the closed form solution.

4.3.1 Additional Assumptions

- Each target follows a linear Gaussian dynamic model and the sensor also follows a linear Gaussian measurement model.The models can be given as follows:

$$f_{k+1|k}(\mathbf{x}|\mathbf{x}') = \mathcal{N}(\mathbf{x}; F_k\mathbf{x}', Q_k) \tag{4.61}$$

$$f_{k+1}(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\mathbf{z}; H_{k+1}\mathbf{x}, R_{k+1}) \tag{4.62}$$

  where $\mathcal{N}(x; m, P)$ is the Gaussian density function with mean m and co-variance P as defined in equation 2.12,F_k is the state transition matrix,Q_k is the process noise co-variance,H_{k+1} is the observation matrix,R_{k+1} is the observation noise co-variance and as mentioned before the subscripts k and $k+1$ of a quantity denote that the quantities are taken at time steps t_k and t_{k+1}.

- The survival and detection probabilities are state independent, i.e.,

$$p_{S,k}(\mathbf{x}) = p_{S,k} \tag{4.63}$$

$$p_{D,k}(\mathbf{x}) = p_{D,k} \tag{4.64}$$

  This assumption will be reviewed again in chapter 6.

- The birth PHD is assumed to be a Gaussian Mixture i.e.,

$$b_{k+1|k}(\mathbf{x}) = \sum_{i=1}^{J_{b,k+1}} w_{b,k+1}^{(i)} \mathcal{N}(\mathbf{x}; m_{b,k+1}^{(i)}, P_{b,k+1}^{(i)}) \tag{4.65}$$

  where $J_{b,k+1}, w_{b,k+1}^{(i)}, m_{b,k+1}^{(i)}, P_{b,k+1}^{(i)}, i = 1, ..., J_{b,k+1}$ determine the shape of the birth intensity.Intuitively, this means that $m_{b,k+1}^{(i)}$ are peaks of the birth intensity, the weights $w_{b,k+1}^{(i)}$ give the expected number of new targets originating form these locations and the co-varinces $P_{b,k+1}^{(i)}$ determine the spread of the birth intensity in the vicinity of the peaks. This is a fairly good assumption since any distribution can be approximated by a Gaussian Mixture.

4.3.2 GM-PHD Recursion Equations

Given that the initial PHD i.e., $D_{0|0}(\mathbf{x})$ is a Gaussian Mixture of form $D_{0|0}(\mathbf{x}) = \sum_{i=0}^{J_{0|0}} w_{0|0}^{(i)} \mathcal{N}(\mathbf{x}; m_{0|0}^{(i)}, P_{0|0}^{(i)})$ and implementing the above assumptions in equations 4.31 and 4.58,the GM-PHD Recursion equations can be given as follows (refer [1]).

GM-PHD prediction step

If the posterior intensity at time t_k is a Gaussian Mixture of the form:

$$D_{k|k}(\mathbf{x}) = \sum_{j=1}^{J_k} w_k^{(j)} \mathcal{N}(\mathbf{x}; m_k^{(j)}, P_k^{(j)}) \tag{4.66}$$

then,

$$D_{k+1|k}(\mathbf{x}) = D_{S,k+1|k}(\mathbf{x}) + b_{k+1|k}(\mathbf{x}) \tag{4.67}$$

where $b_{k+1|k}(\mathbf{x})$ is given by equation 4.65 and:

$$D_{S,k+1|k}(\mathbf{x}) = p_{S,k+1} \sum_{j=1}^{J_k} w_k^{(j)} \mathcal{N}(\mathbf{x}; m_{S,k+1|k}^{(j)}, P_{S,k+1|k}^{(j)}) \tag{4.68}$$

$$m_{S,k+1|k}^{(j)} = F_k m_k^{(j)} \tag{4.69}$$

$$P_{S,k+1|k}^{(j)} = Q_k + F_k P_k^{(j)} F_k^T \tag{4.70}$$

GM-PHD corrector step

Let $Z_{k+1} = \{\mathbf{z}_1, ..., \mathbf{z}_m\}$ denote the measurement set at $t_k + 1$ and let the predicted intensity at t_{k+1} obtained from the prediction step be given by :

$$D_{k+1|k}(\mathbf{x}) = \sum_{i=1}^{J_{k+1|k}} w_{k+1|k}^{(i)} \mathcal{N}(\mathbf{x}; m_{k+1|k}^{(i)}, P_{k+1|k}^{(i)}) \tag{4.71}$$

then the posterior intensity at t_{k+1} is a Gaussian Mixture which is given by:

$$D_{k+1|k+1}(\mathbf{x}) = (1 - p_{D,k+1}) D_{k+1|k}(\mathbf{x}) + \sum_{\mathbf{z} \in Z_{k+1}} D_{D,k+1}(\mathbf{x}; \mathbf{z}) \tag{4.72}$$

where,

$$D_{D,k+1}(\mathbf{x}; \mathbf{z}) = \sum_{j=1}^{J_{k+1|k}} w_{k+1}^{(j)}(\mathbf{z}) \mathcal{N}(\mathbf{x}; m_{k+1|k+1}^{(j)}(\mathbf{z}), P_{k+1|k+1}^{(j)}) \tag{4.73}$$

$$w_{k+1}^{(j)}(\mathbf{z}) = \frac{p_{D,k+1} w_{k+1|k}^{(j)} q_{k+1}^{(j)}(\mathbf{z})}{\kappa_{k+1}(\mathbf{z}) + p_{D,k+1} \sum_{l=1}^{J_{k+1|k}} w_{k+1|k}^{(l)} q_{k+1}^{(l)}(\mathbf{z})} \tag{4.74}$$

$$q_{k+1}^{(j)}(\mathbf{z}) = \mathcal{N}(\mathbf{z}; H_{k+1} m_{k+1|k}^{(j)}, R_{k+1} + H_{k+1} P_{k+1|k}^{(j)} H_{k+1}^T) \tag{4.75}$$

$$m_{k+1|k+1}^{(j)} = m_{k+1|k}^{(j)} + K_{k+1}^{(j)} (\mathbf{z} - H_{k+1} m_{k+1|k}^{(j)}) \tag{4.76}$$

$$P_{k+1|k+1}^{(j)} = [I - K_{k+1}^{(j)} H_{k+1}] P_{k+1|k}^{(j)} \tag{4.77}$$

$$K_{k+1}^{(j)} = P_{k+1|k}^{(j)} H_{k+1}^T (H_{k+1} P_{k+1|k}^{(j)} H_{k+1}^T + R_{k+1})^{-1} \tag{4.78}$$

Summary

From the prediction and the corrector steps it is clear that if the initial prior intensity $D_{0|0}(\mathbf{x})$ is a Gaussian Mixture and the assumptions in section 4.3.1 are made

, then all the subsequent predicted intensities $D_{k+1|k}(\mathbf{x})$ and the posterior intensities $D_{k+1|k+1}(\mathbf{x})$ are Gaussian Mixtures.Equation 4.67 gives a closed-form expressions for evaluating the means,weights and co-variances for $D_{k+1|k}(\mathbf{x})$ from the Gaussian Mixture components of $D_k(\mathbf{x})$ and equations 4.97-4.77 gives the closed-form expressions for the weights,means and co-variances of $D_{k+1|k+1}(\mathbf{x})$ from the measurement set Z_{k+1} at time t_{k+1} and $D_{k+1|k}(\mathbf{x})$.

It can be seen that the predicted intensity $D_{k+1|k}(\mathbf{x})$ is sum of two terms where the first term($D_{S,k+1|k}(\mathbf{x})$) accounts for the surviving targets and the second term($b_{k+1|k}(\mathbf{x})$) accounts for the new targets entering the scene.Similarly, the updated posterior intensity consists of two terms where the first term accounts for missed detections of targets($(1-p_{D,k+1})D_{k|k-1}(\mathbf{x})$) and and the second term ($D_{D,k+1}(\mathbf{x};\mathbf{z})$) accounts for the detected objects.An important thing to note here is that the weights of the Gaussian Components of the detection term($D_{D,k+1}(\mathbf{x};\mathbf{z})$) are calculated by taking into account the association of every measurement with every predicted term.Consider the weight term $w_{k+1}^{(j)}(\mathbf{z})$ in equation 4.97.The numerator calculates the probability that the target at $m_{k+1|k^{(j)}}$ creates the measurement $\mathbf{z}$ and weighs it against the sum of all such probabilities in the denominator which also includes the clutter term.Hence the weight($w_{k+1}^{(j)}(\mathbf{z})$) will be higher if the predicted intensity mean($m_{k+1|k^{(j)}}$) creates the measurement $\mathbf{z}$.It can be seen that the recursion for means and co-variances of $D_{S,k+1|k}(\mathbf{x})$ is the Kalman prediction step and the recursion for $D_{D,k+1}(\mathbf{x};\mathbf{z})$ is a Kalman update.

The total number of targets at any time can be obtained by integrating the PHD over the whole state space $\mathfrak{X}_0$.In the case of GM-PHD filter the expected number of targets $\hat{N}_{k+1|k}$ and $\hat{N}_{k+1|k+1}$ by summing up the appropriate weights of the Gaussian Mixture i.e.,

$$\hat{N}_{k+1|k} = p_{S,k+1}\hat{N}_k + \sum_{j=1}^{J_k} w_k^{(j)} \tag{4.79}$$

$$\hat{N}_{k+1|k+1} = (1-p_{D,k+1})\hat{N}_{k+1|k} + \sum_{\mathbf{z}\in Z_{k+1}} \sum_{j=1}^{J_{k+1|k}} w_{k+1}^{(j)}(\mathbf{z}) \tag{4.80}$$

These equations can be obtained by observing that $\int_{\mathfrak{X}_0} \mathcal{N}(\mathbf{x}; m, P)d\mathbf{x} = 1$.As was discussed before , the expected number of predicted targets can be interpreted as the sum of the expected number of surviving targets and new targets(birth).Similarly after the GM-PHD update, the expected number of targets are the sum of non-detected targets and the detected targets.

Implementation Issues

The GM-PHD filter becomes computationally expensive as time progresses. Consider the term $\sum_{\mathbf{z}\in Z_{k+1}} D_{D,k+1}(\mathbf{x}; \mathbf{z})$ of the update step.For every measurement, this term introduces $J_{k+1|k}$(see equation 4.96) new terms or new Gaussian components.Hence, at time t_{k+1},GM-PHD filter requires $(J_k + J_{b,k+1})(1 + |Z_{k+1}|)$ terms to represent $D_{k+1|k+1}(\mathbf{x})$ which increases as time progresses. In order to avoid this, procedures known as *pruning,merging* and *capping* are used.Pruning approximates a Gaussian Mixture by removing the terms with lowest weights i.e., a threshold for the weight is set and the Gaussian components having weights below this threshold are removed from the mixture.Capping is a similar process but instead of having a threshold on the magnitude of a weight, a threshold on the number of components is set.For example only the Gaussian components having the 100 highest weights can be retained at every time step.Merging is the process of combining two or more Gaussian Components if they are sufficiently close to each other.The following merging approach is used in [1], [39], [40]. Let $(w_1, m_1, P_1), (w_2, m_2, P_2)$ be the weights,means and co-varinaces of two Gaussian Components.Consider the distance metric $d_{1,2}$,

$$d_{1,2} = (m_1 - m_2)^T P_1^{-1}(m_1 - m_2) \tag{4.81}$$

This metric is known as the mahanalobis distance(refer [41]).This distance gives the distance between a point and a Gaussian distribution taking into account the co-variance of the distribution i.e., Consider a point which is far from the mean.It can be "close" to this distribution if the co-variance is high.The mahanalobis distance

takes this property into account.If $d_{1,2}$ is less than a specific threshold, then the components are merged into a single component(w, m, P),where:

$$w = w_1 + w_2 \tag{4.82}$$

$$m = \frac{w_1 m_1 + w_2 m_2}{w_1 + w_2} \tag{4.83}$$

$$P = \frac{w_1 P_1 + w_2 P_2}{w_1 + w_2} + (m_1 - m_2)(m_1 - m_2)^T \tag{4.84}$$

A similar procedure can be applied when multiple Gaussian Components are to be merged(refer [1]).

Extension to Non-linear Gaussian Models

The GM-PHD filter presented in the previous sections assume that the target and birth models are linear. This section extends the GM-PHD filter to the non-linear target models.A detailed proof can be found in [1].
Let the state and the measurement processes be given by the following:

$$\mathbf{x}_{k+1} = \Phi(\mathbf{x}_k, \nu_k) \tag{4.85}$$

$$\mathbf{z}_{k+1} = h_{k+1}(\mathbf{x}_{k+1}, \epsilon_{k+1}) \tag{4.86}$$

where Φ_k and h_{k+1} are non-linear functions and ν_{k+1}(Process Noise),ϵ_{k+1}(measurement noise) are zero-mean Gaussian random variables with co-variances Q_k and R_{k+1} respectively.Because of this non-linearity the predicted and posterior intensities will no longer be Gaussian Mixtures even when the initial Gaussian Mixture($D_{0|0}(\mathbf{x})$) is a Gaussian Mixture.In order to get a closed form solution in this case certain approximations have to be made.The posterior intensity is approximated as a Gaussian by locally linearizing the motion(Φ_k) and measurement models h_{k+1} and then applying the Kalman recursions to it(as was seen in section 2.3).Using these approximations the EK-PHD filter can be developed and the predicted and corrector step can be given as follows(detailed proof in [1]).

EK-PHD prediction step

If the posterior intensity at time t_k is a Gaussian Mixture of the form:

$$D_{k|k}(\mathbf{x}) = \sum_{j=1}^{J_k} w_k^{(j)} \mathcal{N}(\mathbf{x}; m_k^{(j)}, P_k^{(j)}) \tag{4.87}$$

then,

$$D_{k+1|k}(\mathbf{x}) = D_{S,k+1|k}(\mathbf{x}) + b_{k+1|k}(\mathbf{x}) \tag{4.88}$$

where $b_{k+1|k}(\mathbf{x})$ is given by equation 4.65 and:

$$D_{S,k+1|k}(\mathbf{x}) = p_{S,k+1} \sum_{j=1}^{J_k} w_k^{(j)} \mathcal{N}(\mathbf{x}; m_{S,k+1|k}^{(j)}, P_{S,k+1|k}^{(j)}) \tag{4.89}$$

$$m_{S,k+1|k}^{(j)} = \Phi_k(m_k^{(j)}, 0) \tag{4.90}$$

$$P_{S,k+1|k}^{(j)} = G_k^{(j)} Q_k [G_k^{(j)}]^T + F_k^{(j)} P_k^{(j)} [F_k^{(j)}]^T \tag{4.91}$$

$$F_k^{(j)} = \left. \frac{\partial \Phi_k(\mathbf{x}_k^{(j)}, 0)}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = m_k^{(j)}} \tag{4.92}$$

$$G_k^{(j)} = \left. \frac{\partial \Phi_k(m_k^{(j)}, \nu_k)}{\partial \nu_k} \right|_{\nu_k = 0} \tag{4.93}$$

GM-PHD corrector step

Let $Z_{k+1} = \{\mathbf{z}_1, ..., \mathbf{z}_m\}$ denote the measurement set at $t_k + 1$ and let the predicted intensity at t_{k+1} obtained from the prediction step be given by :

$$D_{k+1|k}(\mathbf{x}) = \sum_{i=1}^{J_{k+1|k}} w_{k+1|k}^{(i)} \mathcal{N}(\mathbf{x}; m_{k+1|k}^{(i)}, P_{k+1|k}^{(i)}) \tag{4.94}$$

then the posterior intensity at t_{k+1} is a Gaussian Mixture which is given by:

$$D_{k+1|k+1}(\mathbf{x}) = (1 - p_{D,k+1}) D_{k|k-1}(\mathbf{x}) + \sum_{\mathbf{z} \in Z_{k+1}} D_{D,k+1}(\mathbf{x}; \mathbf{z}) \tag{4.95}$$

where,

$$D_{D,k+1}(\mathbf{x};\mathbf{z}) = \sum_{j=1}^{J_{k+1|k}} w_{k+1}^{(j)}(\mathbf{z})\mathcal{N}(\mathbf{x}; m_{k+1|k+1}^{(j)}(\mathbf{z}), P_{k+1|k+1}^{(j)}) \tag{4.96}$$

$$w_{k+1}^{(j)}(\mathbf{z}) = \frac{p_{D,k+1} w_{k+1|k}^{(j)} q_{k+1}^{(j)}(\mathbf{z})}{\kappa_{k+1}(\mathbf{z}) + p_{D,k+1}\sum_{l=1}^{J_{k+1|k}} w_{k+1|k}^{(l)} q_{k+1}^{(l)}(\mathbf{z})} \tag{4.97}$$

$$q_{k+1}^{(j)}(\mathbf{z}) = \mathcal{N}(\mathbf{z}; \eta_{k+1}^{(j)}, S_{k+1}^{(j)}) \tag{4.98}$$

$$m_{k+1|k+1}^{(j)} = m_{k+1|k}^{(j)} + K_{k+1}^{(j)}(\mathbf{z} - \eta_{k+1}^{(j)}) \tag{4.99}$$

$$\eta_{k+1}^{(j)} = h_{k+1}(m_{k+1|k}^{(j)}, 0) \tag{4.100}$$

$$S_{k+1}^{(j)} = U_{k+1}^{(j)} R_{k+1} [U_{k+1}^{(j)}]^T + H_{k+1}^{(j)} P_{k+1|k}^{(j)} [H_{k+1}^{(j)}]^T \tag{4.101}$$

$$P_{k+1|k+1}^{(j)} = [I - K_{k+1}^{(j)} H_{k+1}^{(j)}] P_{k+1|k}^{(j)} \tag{4.102}$$

$$K_{k+1}^{(j)} = P_{k+1|k}^{(j)} [H_{k+1}^{(j)}]^T [S_{k+1}^{(j)}]^{-1} \tag{4.103}$$

$$H_{k+1}^{(j)} = \left.\frac{\partial h_{k+1}(\mathbf{x}_{k+1}, 0)}{\partial \mathbf{x}_{k+1}}\right|_{\mathbf{x}_{k+1}=m_{k+1|k}^{(j)}} \tag{4.104}$$

$$U_{k+1}^{(j)} = \left.\frac{\partial h_{k+1}(m_{k+1|k}^{(j)}, \epsilon_{k+1})}{\partial \epsilon_{k+1}}\right|_{\epsilon_{k+1}=0} \tag{4.105}$$

The target model used in this thesis is the two-body motion which is non-linear(described in later sections).Hence, the EK-PHD filter will be used for this thesis.So, implementing the EK-PHD filter will result in a Gaussian Mixture at ever time step.The following section describes the method used to estimate the target states form the Gaussian Mixture.

State Estimation

Since the posterior intensity ($D_{k+1|k+1}(\mathbf{x})$) is a Gaussian Mixture, the means of the Gaussian components are the multi-target state estimates provided the components are well separated(this is ensured by the pruning process).From equation 4.97,the expected number of targets can be evaluated and selecting the $\hat{N}_{k+1|k+1}$ "peaks" from the Gaussian Mixture($D_{k+1|k+1}(\mathbf{x})$) one can get the state estimate at any time. But doing this is undesirable since this procedure would estimate a state corresponding

to a Gaussian with weak weight.Consider a component which has a low weight.The magnitude of the peak can still be high for this component since the height of each peak depends on the weight and the co-variance.Then following the procedure described above results in a bad state estimate since the expected number of targets for that particular component is less (since weight is less).In order to avoid this, only the means of the Gaussian components which have weights greater than a specific threshold(say 0.5) are used as state estimates.A detailed explanation can be found in [1].

The next section describes the target and the measurement models used in this research to generate the results.

5. SIMULATION MODELS

This chapter details the target motion model and the single sensor measurement model which will be used in the application of the EK-PHD filter. Before the description of the motion and measurement model, some important coordinate systems are described first. The description given below is obtained from [42].

5.1 Coordinate Systems

5.1.1 Geocentric Equitorial System

The defining characteristics of this frame are listed below:

- Origin:Center of the Earth
- Fundamental Plane(Plane defined by the $\mathbf{x}$ and $\mathbf{y}$ axes):Plane containing the Equator at a fixed equinox.
- Reference direction($\mathbf{x}$ axis direction):Vernal Equinox at a fixed direction.
- Right Handed Orthogonal coordinate system
- Coordinates : α(right-ascension), δ(declination), r(radial distance from the origin or $\mathbf{x}, \mathbf{y}, \mathbf{z}$
- The position vector can then be described as $\vec{r} = \mathbf{x}\hat{e}_{\mathbf{x}} + \mathbf{y}\hat{e}_{\mathbf{y}} + \mathbf{z}\hat{e}_{\mathbf{z}}$ or $\vec{r} = r\cos(\delta)\cos(\alpha)\hat{e}_{\mathbf{x}} + r\cos(\delta)\sin(\alpha)\hat{e}_{\mathbf{y}} + r\sin(\delta)\hat{e}_{\mathbf{z}}$

The following figure illustrates the Geocentric Coordinate System.

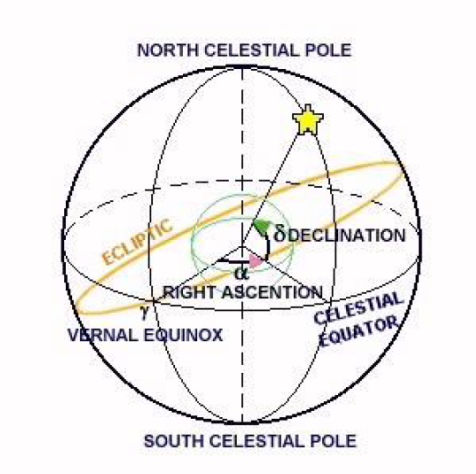


Figure 5.1. Geocentric Coordinate System (Image taken from [42])

The coordinates (r, α, δ) and $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ are related as follows:

$$x = r\cos(\delta)\cos(\alpha) \tag{5.1}$$

$$y = r\cos(\delta)\sin(\alpha) \tag{5.2}$$

$$z = r\sin(\delta) \tag{5.3}$$

Then,

$$r = \sqrt{\mathbf{x}^2 + \mathbf{y}^2 + \mathbf{z}^2} \tag{5.4}$$

$$\tilde{r} = \sqrt{\mathbf{x}^2 + \mathbf{y}^2} \tag{5.5}$$

$$\delta = \begin{cases} \frac{\pi}{2}, 0, -\frac{\pi}{2} & \text{for } \tilde{r} = 0 \text{ and } \mathbf{z} > 0, z = 0, z < 0 \\ \arctan(\frac{\mathbf{z}}{\tilde{r}}) & \text{for } \tilde{r} \neq 0 \end{cases} \tag{5.6}$$

$$\phi = \arctan(\frac{\mathbf{y}}{|\mathbf{x}|}) \tag{5.7}$$

$$\alpha = \begin{cases} 0 & \text{for } \mathbf{x} = 0 \text{ and } \mathbf{y} = 0 \\ \phi & \text{for } \mathbf{x} \geq 0 \text{ and } \mathbf{y} \geq 0 \\ 2\pi + \phi & \text{for } \mathbf{x} \geq 0 \text{ and } \mathbf{y} \leq 0 \\ \pi - \phi & \text{for } \mathbf{x} \leq 0 \end{cases} \tag{5.8}$$

α denotes the in-plane angle which is angle made by the projection of the position vector with the **x**-axis and δ is the angle made by the position vector with the **z**-axis. The Geocentric Equitorial System is a good approximation of a inertial system and hence it will be used to define the motion model of the targets.

5.1.2 Topocentric Equitorial System

Most of the observations are made from the surface of the Earth. Hence, it is necessary to define a coordinate system with the observer's location (also known as topocenter) as the origin. Topocentric Coordinate system is obtained by shifting the Geocentric Coordinate system to the topocenter. The main characteristics of this system are:

- Origin:Topocenter(time-dependent)
- Fundamental Plane(Plane defined by the **x** and **y** axes):Plane containing the Equator at a fixed equinox.
- Reference direction(**x** axis direction):Vernal Equinox at a fixed direction.
- Right Handed Orthogonal coordinate system
- Coordinates : α'(right-ascension), δ'(declination), ρ(radial distance from the topocenter) where $-\frac{\pi}{2} \leq \delta' \leq \frac{\pi}{2}, 0 \leq \alpha' \leq 2\pi, \rho \geq 0$ or $\mathbf{x}', \mathbf{y}', \mathbf{z}'$.

The origin of this system is time dependent and hence it's tracked w.r.t the Geocentric Equitorial System. Sidereal time (θ) is defined as the right-ascension of the topocenter in the Geocentric frame. Hence, observers on the same longitude will have the same sidereal time. Hour angle (τ) is defined as the difference between the sidereal time and the right-ascension of object in the Geocentric frame ($\tau = \theta - \alpha$). This quantity will be useful in defining the local-horizon coordinate system which is discussed in the next section. The relation between the angles(α', δ') and the coordinates $\mathbf{x}', \mathbf{y}', \mathbf{z}'$ are identical to the ones given in previous section except the the coordinates $\mathbf{x}', \mathbf{y}', \mathbf{z}'$ are taken w.r.t to the topocentric equitorial frame.

5.1.3 Topocentric Local Horizon Coordinate System

This coordinate system can be used to determine if the target can be viewed from a topocenter. Important characteristcs are listed below:

- Origin:Topocenter(time-dependent)
- Fundamental Plane(Plane defined by the **x** and **y** axes):Local Horizon
- Reference direction(**x** axis direction):South
- Left Handed Orthogonal coordinate system
- Coordinates : h(elevation), a(azimuth) where $-\frac{\pi}{2} \leq h \leq \frac{\pi}{2}, 0 \leq a \leq 2\pi$

Let hour angle(τ)=$\theta-\alpha$, sidereal time=θ,geographic latitude of the topocenter(ϕ),topocentric equitorial coordinates of the object=α', δ',Local horizon coordinates:a(azimuth), h(elevation). The relation between the local horizon co-ordinates and the topocentric equitorial co-ordinates is then given by:

$$cos(h)cos(a) = sin(\phi)cos(\delta')cos(\tau) - cos(\phi)sin(\delta') \tag{5.9}$$

$$cos(h)sin(a) = cos(\delta')sin(\tau) \tag{5.10}$$

$$sin(h) = sin(\phi)sin(\delta') + cos(\phi)cos(\delta')cos(\tau) \tag{5.11}$$

If the elevation angle $h \leq 0$, it implies that the object is beneath the local horizon plane which makes the object unobservable. Hence the object is has a chance of being detected only if it's above the local horizon plane i.e $h \geq 0$.

5.2 Target Motion Model

The target motion model is assumed to be the Two-body motion model where one body is the Earth and the other body is the target (debris,satellite). The motion model can be expressed as follows:

$$\ddot{\vec{r}} = -\frac{\mu_{\oplus}\vec{r}}{|r|^3} \tag{5.12}$$

where $\vec{r} = \mathbf{x}\hat{e}_{\mathbf{x}} + \mathbf{y}\hat{e}_{\mathbf{y}} + \mathbf{z}\hat{e}_{\mathbf{z}}, |r| = \sqrt{\mathbf{x}^2 + \mathbf{y}^2 + \mathbf{z}^2}$ is position vector of the object in the Geocentric Coordinate frame(section 5.1.1),$\mu_\oplus$ is Gravitational parameter of Earth . Let the state of the object be denoted by $X = [\mathbf{x},\mathbf{y},\mathbf{z},\dot{\mathbf{x}},\dot{\mathbf{y}},\dot{\mathbf{z}}]$. Therefore, the dimension of the state $n_{\mathbf{x}} = 6$. Then, the dynamic state equation can be given by:

$$\dot{X} = A(X)X \tag{5.13}$$

i.e.,

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \mathbf{y} \\ \mathbf{z} \\ \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \\ \dot{\mathbf{z}} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -\frac{\mu_\oplus \mathbf{x}}{|r|^3} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{\mu_\oplus \mathbf{y}}{|r|^3} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{\mu_\oplus \mathbf{z}}{|r|^3} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \\ \mathbf{z} \\ \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \\ \dot{\mathbf{z}} \end{bmatrix} \tag{5.14}$$

5.3 Sensor Measurement Model

Every measurement is taken in the Topocentric Equitorial Frame and assumed to consist of the angles and the angle rates i.e., $\alpha, \dot{\alpha}, \delta, \dot{\delta}$ where the angles are given as described previously(equations 5.6 and 5.8) and the angle rates are given by:

$$\dot{\alpha} = \frac{\mathbf{x}\dot{\mathbf{y}} - \mathbf{y}\dot{\mathbf{x}}}{\tilde{r}^2} \tag{5.15}$$

$$\dot{\delta} = \frac{\tilde{r}}{r^2}\left(\dot{\mathbf{z}} - \frac{\mathbf{z}(\mathbf{x}\dot{\mathbf{x}} + \mathbf{y}\dot{\mathbf{y}})}{\tilde{r}^2}\right) \tag{5.16}$$

Therefore a measurement $Z = [\alpha, \dot{\alpha}, \delta, \dot{\delta}]$ and the dimension of a measurement $n_{\mathbf{z}} = 4$. Note that here $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dot{\mathbf{x}}, \dot{\mathbf{y}}, \dot{\mathbf{z}}$ are taken w.r.t the Topocentric Equitorial Frame. More details regarding the topocenter and the number of measurements are given in the results chapter(Chapter 7).

5.3.1 Field of View(FOV)

An optical measurement(angles and angle rates) is usually obtained with the help of a telescope and hence there is always a field of view associated with it. For exam-

ple,a satellite or a target cannot be observed if it is below the local horizon of the observer. Also, the FOV is modeled differently for objects in GEO(Geostationary Equitorial Orbit) and LEO(Low Earth Orbit). For the case of GEO objects, it is assumed that the sensor has a FOV of $x^\circ \times y^\circ$. This means that the objects will only be in the FOV of the sensor if they satisfy the following conditions:

$$\alpha_{obs} - \frac{x^\circ}{2} \leq \alpha_{obj} \leq \alpha_{obs} + \frac{x^\circ}{2} \tag{5.17}$$

$$\delta_{obs} - \frac{y^\circ}{2} \leq \delta_{obj} \leq \delta_{obs} + \frac{y^\circ}{2} \tag{5.18}$$

where $\alpha_{obs}, \delta_{obs}$ are the location of the observer in the Geocentric Equitorial Coordinate frame and $\alpha_{obj}, \delta_{obj}$ are the coordinates of the object in Topocentric Equitorial Coordinate system.

The same model cannot be used for the LEO objects since these move relatively fast and are difficult to capture in a observer's FOV for a long time whereas since the GEO objects move with a particular location on Earth, the FOV model described above is appropriate. Hence, in case of the LEO objects, the observing telescope is assumed to follow a particular object and the FOV is then assumed to be a window around that object with a FOV of ; $x^\circ \times y^\circ$ i.e., an object is observed if it satisfies

$$\alpha_{ref} - \frac{x^\circ}{2} \leq \alpha_{obj} \leq \alpha_{ref} + \frac{x^\circ}{2} \tag{5.19}$$

$$\delta_{ref} - \frac{y^\circ}{2} \leq \delta_{obj} \leq \delta_{ref} + \frac{y^\circ}{2} \tag{5.20}$$

where $\alpha_{ref}, \delta_{ref}$ are the location of the reference object which the observer follows in the Topocentric Equitorial Coordinate frame and $\alpha_{obj}, \delta_{obj}$ are the coordinates of the observed objects in Topocentric Equitorial Coordinate system. Also, an object is only observed if it is above the local horizon. Hence this condition should also be checked in the case of LEO objects. This can be verified by the elevation coordinate of the Topocentric Local Horizon Coordinate system i.e., an object can be observed if $h > 0$ a certain threshold h_0(taking the obstacles(buildings etc) into account). Also, the FOV can be chenged by focussing another object in it's present FOV. The following figure depicts this process.

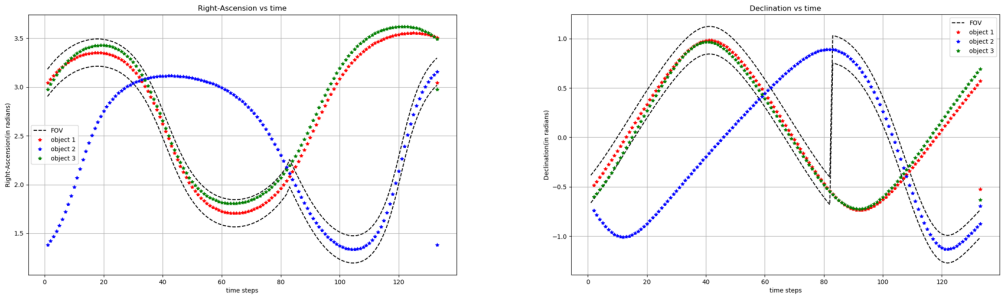


Figure 5.2. FOV movement when observing LEO objects

In the above figure,the black dotted lines represent the movement of FOV with the objects with FOV. It can also be seen that the FOV starts following object 3 at around time step 80.

5.4 Birth Model

As discussed in the previous section, the GM-PHD filter assumes that the birth intensity $b_{k+1|k}(\mathbf{x})$ is described by a Gaussian Mixture. If the state of the target which enters the scene at any time instant is known with a certain uncertainty then the birth Gaussian Mixture is defined in a straightforward manner. But in the space problem it is difficult to identify when and where a object will enter the scene. Also, since the GM-PHD filter doesn't identify the objects it tracks, it makes it difficult to identify new objects.

The only information present on the new and existing targets comes from the measurements. Hence,the birth model which is defined in the next section is based on the measurement. The concept of Admissible Region(AR) is used to obtain the Birth Gaussian Mixture. AR defines all the regions of possible state space corresponding to Earth-captured orbits given a measurement. These regions are then approximated by a Gaussian Mixture to obtain the Birth Model. The next sections details these steps.

5.4.1 Admissible Region(AR)

The Admissible Region approach was developed by Milani et al. [43] and Tommei et al. [44]. Given a measurement of angles(α, δ) and angle-rates($\dot{\alpha}, \dot{\delta}$), AR gives a set of possible range(ρ), range-rate($\dot{\rho}$) such that the resultant orbit is an Earth Captured orbit. Note that the six variables are necessary to define the state of a target. It can be either [**x,y,z**, $\dot{\mathbf{x}}, \dot{\mathbf{y}}, \dot{\mathbf{z}}$] or $[\rho, \dot{\rho}, \alpha, \dot{\alpha}, \delta, \dot{\delta}]$. Given a measurement(or a Admissible region) and no additional region no single combination of range and range-rate can be said to be more probable than the others. Hence,in a probabilistic sense it can be said that the range,range-rate space or the AR can be described by a 2-D Uniform Distribution. The final step is to approximate this uniform distribution by a Gaussian Mixture to obtain the birth intensity. This process was developed in [45]. The next section reviews the concept of admissible region(A detailed discussion can be found in [44], [43]).

Review of Admissible Region

Let the measurement be given by $Z = [\alpha, \dot{\alpha}, \delta, \dot{\delta}]$. The Two-body energy can be given by:

$$\mathcal{E} = \frac{\|\dot{\vec{\mathbf{r}}}\|^2}{2} - \frac{\mu_{\oplus}}{\|\vec{\mathbf{r}}\|} \tag{5.21}$$

where $\vec{\mathbf{r}}$ is the inertial position of the target(position vector w.r.t the Geocentric Frame) and $\dot{\vec{\mathbf{r}}}$ is the inertial velocity of the target(velocity vector w.r.t the Geocentric Frame). Since the measurements are made in the topocentric frame,$\vec{\mathbf{r}}$ and $\dot{\vec{\mathbf{r}}}$ are given as:

$$\vec{\mathbf{r}} = \vec{\mathbf{q}} + \vec{\boldsymbol{\rho}} \tag{5.22}$$

$$\dot{\vec{\mathbf{r}}} = \dot{\vec{\mathbf{q}}} + \dot{\vec{\boldsymbol{\rho}}} \tag{5.23}$$

where $\vec{\mathbf{q}}, \dot{\vec{\mathbf{q}}}$ are the inertial position and velocity of the observer in Geocentric Equitorial Frame and $\vec{\boldsymbol{\rho}}, \dot{\vec{\boldsymbol{\rho}}}$ are the inertial position and velocity of the target in Topocentric

Equitorial Frame. Then, $\rho = \|\vec{\boldsymbol{\rho}}\|, \alpha, \delta$ are the spherical co-ordinates of the target with topocenter as the center. Then $\vec{\boldsymbol{\rho}}$ and range-rate($\dot{\vec{\boldsymbol{\rho}}}$) are given as:

$$\vec{\boldsymbol{\rho}} = \rho \mathbf{u}_\rho \tag{5.24}$$

$$\dot{\vec{\boldsymbol{\rho}}} = \dot{\rho}\mathbf{u}_\rho + \rho\dot{\alpha}\mathbf{u}_\alpha + \rho\dot{\delta}\mathbf{u}_\delta \tag{5.25}$$

where $\mathbf{u}_\rho, \mathbf{u}_\alpha, \mathbf{u}_\delta$ are unit vectors given by:

$$\mathbf{u}_\rho = \begin{bmatrix} \cos(\alpha)\cos(\delta) \\ \sin(\alpha)\cos(\delta) \\ \sin(\delta) \end{bmatrix}, \mathbf{u}_\alpha = \begin{bmatrix} -\sin(\alpha)\cos(\delta) \\ \cos(\alpha)\cos(\delta) \\ 0 \end{bmatrix}, \mathbf{u}_\alpha = \begin{bmatrix} -\cos(\alpha)\sin(\delta) \\ -\sin(\alpha)\sin(\delta) \\ 0 \end{bmatrix}$$

Let, $w_0 = \|\vec{\mathbf{q}}\|^2, w_1 = 2(\vec{\mathbf{q}}.\mathbf{u}_\rho), w_2 = \dot{\alpha}^2\cos(\delta)^2 + \dot{\delta}^2, w_3 = 2\dot{\alpha}(\dot{\mathbf{q}}.\mathbf{u}_\alpha) + 2\dot{\delta}(\dot{\mathbf{q}}.\mathbf{u}_\delta)$, $w_4 = \|\dot{\vec{\mathbf{q}}}\|^2,$ and $w_5 = 2(\vec{q}.\mathbf{u}_\rho)$.Then $\|\vec{\mathbf{r}}\|, \|\dot{\vec{\mathbf{r}}}\|$ can be written as:

$$\|\vec{\mathbf{r}}\| = \rho^2 + w_5\rho + w_0 \tag{5.26}$$

$$\|\dot{\vec{\mathbf{r}}}\| = \dot{\rho}^2 + w_1\dot{\rho} + w_2\rho^2 + w_3\rho + w_4 \tag{5.27}$$

Then substituting the above equations in the Energy equation(equation 5.21),it can be seen that:

$$\dot{\rho}^2 + w_1\dot{\rho} - \mathcal{F}(\rho) - 2\mathcal{E} = 0 \tag{5.28}$$

where,

$$\mathcal{F}(\rho) = w_2\rho^2 + w_3\rho + w_4 - \frac{2\mu_\oplus}{\sqrt{\rho^2 + w_5\rho + w_0}} \tag{5.29}$$

Therefore given a value of the range(ρ), the two(equation 5.28 is quadratic) possible values of the range-rate($\dot{\rho}$) are given as:

$$\dot{\rho} = -\frac{w_1}{2} \pm \sqrt{\left(\frac{w_1}{2}\right)^2 - \mathcal{F}(\rho) + 2\mathcal{E}} \tag{5.30}$$

Using the above equation, the AR can be plotted given a measurement and observer location. As an example, the following measurement and observer information will be used throughout this section.

$$Z = [\alpha, \dot{\alpha}, \delta, \dot{\delta}] = [10^\circ, 15^\circ/hr, -2^\circ, 3^\circ/hr] \tag{5.31}$$

The topocenter(observer) has the following location(λ,latitude(declination) and ϕ,longitude(right-ascension));

$$\lambda = 0°, \phi = 30° \tag{5.32}$$

$\vec{\mathbf{q}}, \dot{\vec{\mathbf{q}}}$ are then given by:

$$\vec{\mathbf{q}} = 5523.628\hat{e}_{\mathbf{x}} + 3189.07\hat{e}_{\mathbf{z}} \quad \text{km} \tag{5.33}$$

$$\dot{\vec{\mathbf{q}}} = \vec{w}_{\oplus} \times \vec{\mathbf{q}} \tag{5.34}$$

where $\vec{w}_{\oplus}$ is the angular velocity vector of the Earth. The AR is then plotted as follows:

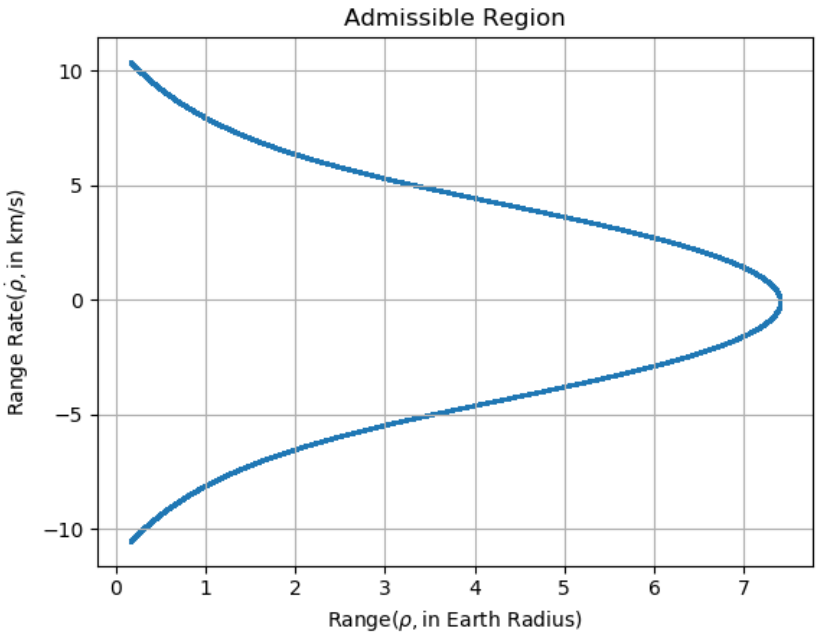


Figure 5.3. Admissible Region

The region inside the blue boundary represents the set of all possible $\rho, \dot{\rho}$ pairs such that the resultant orbit is Earth Captured. From the figure it can be seen that after $\rho = 7.5R_{\oplus}$, ($R_{\oplus}$ is radius of the Earth), there are no Earth captured orbits for this particular measurement i.e., this results in Equation 5.28 having complex roots.

As was mentioned before, under no additional information any point in this region

cannot be assumed to be more probable than the other points. Hence, this region can be aptly represented to have a uniform distribution. The next section describes the procedure to approximate a uniform distribution by a Gaussian Mixture.

5.4.2 Gaussian Mixture Approximation of an Admissible Region

The following details the procedure (developed by authors of [45]) to approximate a uni-variate uniform distribution by a Gaussian mixture. This concept is applied to the 2D case.

Gaussian Mixture Approximation of a Uni-variate Uniform Distribution

A uni-variate uniform distribution in the interval $[a, b]$ can be described as follows:

$$p(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b] \\ 0 & \text{otherwise} \end{cases} \tag{5.35}$$

This distribution is to be approximated by a GM of the form:

$$q(x) = \sum_{i=1}^{L} \alpha_i \mathcal{N}(x; m_i, P_i) \tag{5.36}$$

where L is the number of Gaussian components and α_i, m_i, P_i are the weights,means and co-variances of the Gaussian components respectively. The optimal approximation is obtained by minimizing the following integral (also known as L_2 norm):

$$L_2[p][q] = \int_{\mathcal{R}} (p(x) - q(x))^2 dx \tag{5.37}$$

where $\mathcal{R}$ is the real number space. The optimisation problem can be then stated as:

$$\min L_2[p][q] \text{ subject to } \alpha_i \geq 0 \quad \forall i \text{ and } \sum_{i=1}^{L} \alpha_i = 1 \tag{5.38}$$

The last constraint comes from the requirement that the probability over the entire sample space(here it is $\mathcal{R}$) should be 1. The L_2 norm can be further simplied to the following expression(for detailed proof,refer [45]):

$$L_2[p][q] = \frac{1}{b-a} + \sum_{i=1}^{L}\sum_{j=1}^{L} \alpha_i \alpha_j \Gamma(m_i, m_j, P_i, P_j) - \frac{1}{b-a}\sum_{i=1}^{L} \alpha_i \left[erf\left\{ \frac{b-m_i}{\sqrt{2P_i}} \right\} - erf\left\{ \frac{a-m_i}{\sqrt{2P_i}} \right\} \right] \tag{5.39}$$

where,

$$\Gamma(m_i, m_j, P_i, P_j) = |2\pi(P_i + P_j)^{-1/2}| \exp\left(-\frac{1}{2}(m_i - m_j)^T (P_i + P_j)^{-1} (m_i - m_j) \right) \tag{5.40}$$

where erf is the error function. The optimisation function i.e., the L_2 norm is ill-conditioned due to the existence of large number of local minima. For example, the GM components can be exchanged with no significant change in the L_2 norm. Also, the number of parameters to solve is also large as L increases(need to solve for $3L$ variables(weights,means,co-variances) for L components). Hence, to reduce the number of parameters and make the optimisation problem well conditioned, the following assumptions are made.

- The weights(w_i) are assumed to be equal for all the components.
- The Gaussian Mixture is assumed to be *homoscedastic* i.e., the co-variance(P_i) is assumed to be same for all the components.
- Finally, the means(m_i) are assumed to be evenly distributed across the support of $p(x)$ i.e., in the interval $[a, b]$.

The optimisation function is then reduced to:

$$L_2[p][q] = \frac{1}{b-a} + \frac{\alpha^2}{2\sqrt{\pi}\sigma} \sum_{i=1}^{L}\sum_{j=1}^{L} \exp\left\{ -\frac{1}{4}\left(\frac{m_1 - m_j}{\sigma} \right)^2 \right\} - \frac{\alpha}{b-a}\sum_{i=1}^{L} \alpha_i [erf\{B_i\} - erf\{A_i\}] \tag{5.41}$$

where,

$$A_i = \left(\frac{a-m_i}{\sqrt{2}\sigma} \right) \text{ and } B_i = \left(\frac{b-m_i}{\sqrt{2}\sigma} \right) \tag{5.42}$$

Hence,the optimisation problem is only solved for the parameter σ,given the number of GM components(L). Without loss of generality, it can be assumed that $a = 0, b = 1$. Then a standard library of solutions can be produced for the value of σ by varying L,the number of components. According to the assumptions, the means and weights can be definied as follows:

$$\alpha = \frac{1}{L} \text{ and } m_i = \frac{i}{L+1} \text{ for } i = \{1, ..., L\} \tag{5.43}$$

Substituting these expressions into the L_2 norm(equation 5.41) and solving the optimistion problem results in the optimal value of $\tilde{\sigma}$ for a particular value of L. The following figure shows the $\tilde{\sigma}$ values for $L = 1, ..., 15$.

L	$\tilde{\sigma}$
1	0.3467341
2	0.2902619
3	0.2465658
4	0.2000502
5	0.1531422
6	0.1224991
7	0.1025690
8	0.0884449
9	0.0778498
10	0.0695800
11	0.0629317
12	0.0574633
13	0.0528821
14	0.0489859
15	0.0456302

Figure 5.4. Standard Library(result reproduced from [45])

The above solutions are true for the case when $a = 0, b = 1$. For a general case the Gaussian component parameters are given as follows:

$$\alpha_i = \frac{1}{L}, \quad m_i = a + \frac{(b-a)i}{L+1} \text{ and } \sigma = (b-a)\tilde{\sigma} \quad \forall\, i = \{1, ..., L\} \tag{5.44}$$

Application to Admissible Region

The pdf(probability density function) of the uniform distribution over the AR is over 2 dimensions(range and range-rate). Hence, in order to use the GM approximation strategy developed in the previous section,the 2D approximation must be divided into two 1D approximations. The first step is to approximate the range marginal pdf by a Gaussian Mixture. Let $p_{\rho,\dot{\rho}}(\rho,\dot{\rho})$ denote the the uniform pdf over the Admissible Region. Then the range marginal pdf($p_\rho(\rho)$) is defined as:

$$p_\rho(\rho) = \int_{-\infty}^{\infty} p_{\rho,\dot{\rho}}(\rho, s)ds \tag{5.45}$$

The admissible region boundaries for ρ can be determined directly by a numerical procedure(using equation 5.28). Let the boundaries be a, b with $a < b$. Then, to use the pre-computed standard library a design parameter σ_ρ has to be defined. Using σ_ρ,the number of required Gaussian components can then be found by using the relation $\tilde{\sigma}(b-a) < \sigma_\rho$. The design parameter determines the level of accuracy of the approximation. Lesser the value the more accurate the approximation becomes but the number of components increase. Hence, this trade-off must be decided based on the requirement. The weights,means and co-variances are then given as follows:

$$\alpha_i = \frac{1}{L}, \quad m_i = a + \frac{(b-a)i}{L+1} \text{ and } P_i = ((b-a)\tilde{\sigma})^2 \quad \forall\, i = \{1, ..., L\} \tag{5.46}$$

Note that the GM evaluated above approximates a uniform distribution over the range support set and not the range marginal pdf($p_\rho(\rho)$). However, the range marginal pdf is not required to be a uniform pdf. Figure 5.5 shows the range marginal pdf for the example discussed previously(Figure 5.4). It can be seen that $p_\rho(\rho)$ is not uniform. This can also deduced from Figure 5.4 as the AR is "thicker" at lower values of ρ and becomes "thinner" as ρ increases.

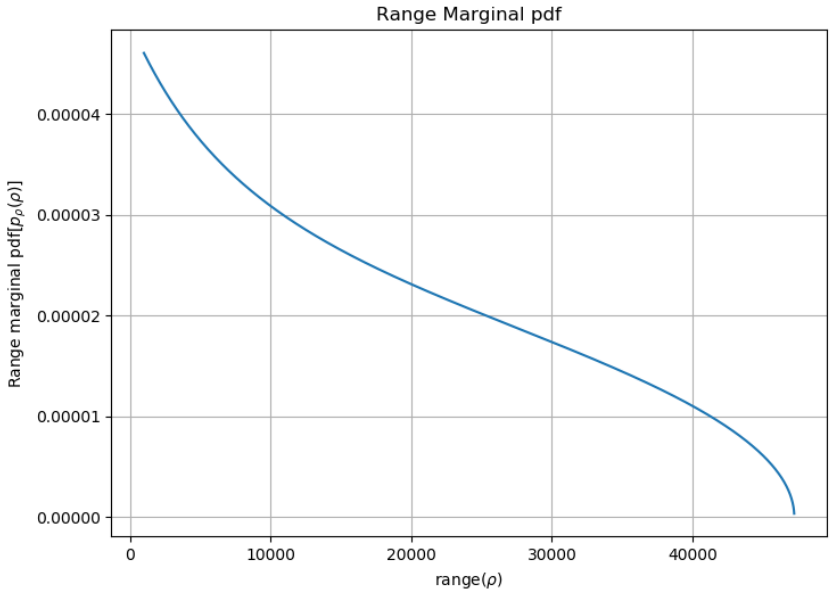


Figure 5.5. Range marginal pdf

To account for this, the weights of the GM components can be changed. The new weights are evaluated using a least squares approach. The steps of this approach are detailed below.

- Firstly,select M points in the range support set i.e., M points between a and b. Let the points be denoted by $\rho_i \; \forall \; i = \{1, ..., M\}$.

- Evaluate $p_\rho(\rho)$ at these points using Equation 5.45. Let the probability be denoted by $p_i = p_\rho(\rho_i) \; \forall \; i = \{1, ..., M\}$.let $\mathbf{p}$ be the vector of the probabilities with the i^{th} element as p_i.

- Then the weights of the Gaussian components are evaluated by solving the following optimisation problem.

$$\min \; J = \|\mathbf{p} - \mathbf{H}\boldsymbol{\alpha}\| \text{ subject to } \alpha_i \geq 0 \; \forall \; i \text{ and } \sum_{i=1}^{L} \alpha_i = 1 \tag{5.47}$$

  where H is a matrix with $\mathbf{H}(i,j) = \mathcal{N}(\rho_i; m_j, P_j)$ and $\boldsymbol{\alpha}$ is the vector of GM weights with the i^{th} element as α_i.

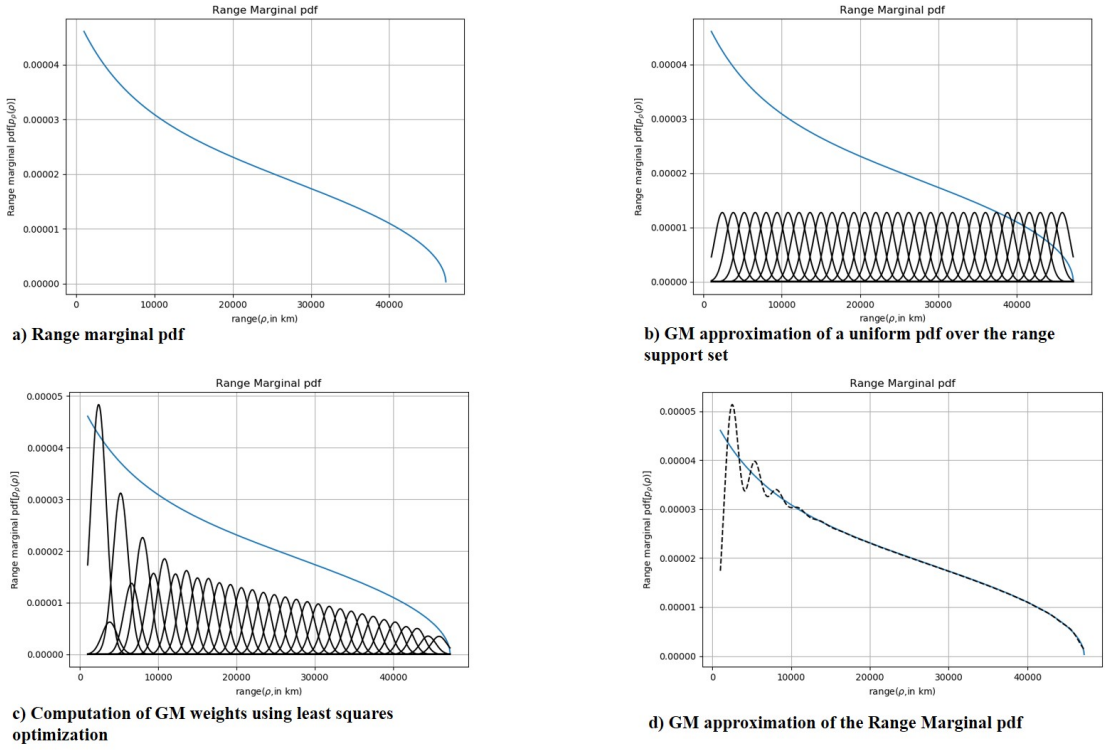


Figure 5.6. Procedure for the GM approximation of Range Marginal pdf

Figure 5.6 illustrates the process described above. The design parameter σ_ρ was taken to be 1000 and the number of components was obtained as 32. From b) and c) it can be seen that the least squares optimization increased the weights of the components near the beginning to account for higher range marginal pdf and similarly the weights become lower as ρ increases because of the decrease in range marginal pdf.

The next step is to move on to the range rate direction. For any component in the GM approximation of the range marginal pdf, no value of range-rate is more or less likely i.e., the uncertainty of the range-rate can be assumed to be uniform. Therefore, a bi-variate Gaussian Mixture marginal pdf can be obtained by applying the previously described GM approximation technique of a univariate uniform pdf in the range-rate

direction for every component of the GM range marginal pdf.

The range-marginal GM pdf is relabeled as follows:

$$p_\rho(\rho) = \sum_{i=1}^{L_\rho} \alpha_{\rho,i} \mathcal{N}(\rho; m_{\rho,i}, P_{\rho,i}) \tag{5.48}$$

The next step is to determine the number of components. Analogous to the range direction, a design parameter $\sigma_{\dot{\rho}}$ is defined which determines the accuracy of the approximation. For each component of the GM range marginal pdf($m_{\rho,i}$), let the support set of range-rate be defined as a_i, b_i with $a_i < b_i$ and the length of the support set as $\Delta_i = b_i - a_i$. Then using the design parameter $\sigma_{\dot{\rho}}$,the support set and the pre-computed library of solutions,the GM approximation for the range-rate uniform pdf(corresponding to the i^{th} component of GM range marginal pdf) is evaluated as:

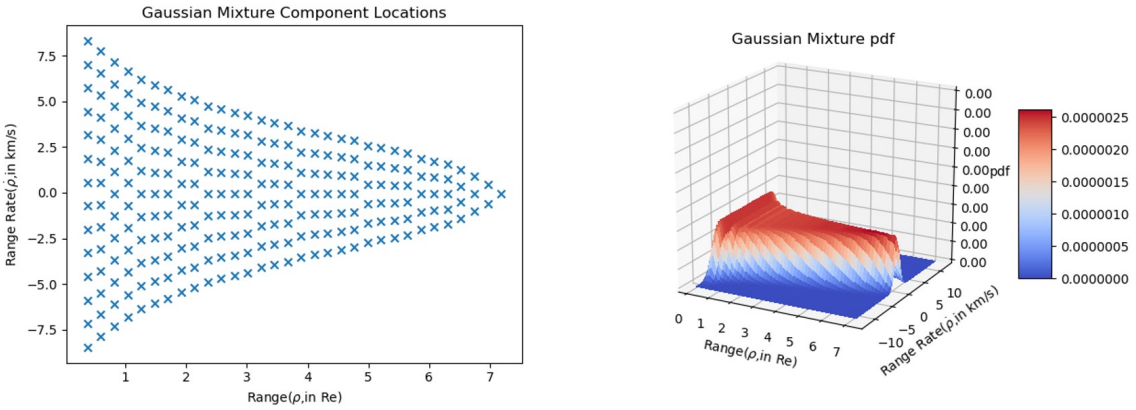
$$\alpha_{\dot{\rho},i} = \frac{1}{L_{\dot{\rho},i}}, m_{\dot{\rho},i} = a_i + \frac{\Delta_i i}{L_{\dot{\rho},i}+1} \text{ and } P_{\dot{\rho},i} = (\Delta_i \tilde{\sigma}^2) \; \forall \; i \in \{1, ..., L_{\dot{\rho},i}\} \tag{5.49}$$

Consider the indexing variable $i_l = \sum_{j=1}^{l-1} L_{\dot{\rho},j}$. Then, given the l^{th} component of the range marginal GM pdf and the corresponding k components of the range-rate GM approximation the components of bi-variate GM approximation are given as:

$$\alpha_{\rho,\dot{\rho},i_l+k} = \alpha_{\rho,l}\alpha_{\dot{\rho},k}, \quad m_{\rho,\dot{\rho},i_l+k} = \begin{bmatrix} m_{\rho,l} \\ m_{\dot{\rho},k} \end{bmatrix} \quad \text{and} \quad P_{\rho,\dot{\rho},i_l+k} = \begin{bmatrix} P_{\rho,l} & 0 \\ 0 & P_{\dot{\rho},k} \end{bmatrix}$$

where $k \in \{1, ..., L_{\dot{\rho},l}\}$. The above procedure is applied for every component of the range marginal GM pdf. The GM approximation of the Admissible Region is then given by:

$$p_{\rho,\dot{\rho}}(\rho, \dot{\rho}) \approx \sum_{i=1}^{L} \alpha_{\rho,\dot{\rho},i} \mathcal{N}(\rho, \dot{\rho}; m_{\rho,\dot{\rho},i}, P_{\rho,\dot{\rho},i}) \tag{5.50}$$



a) Location of GM components

b) pdf of the Gaussian Mixture

Figure 5.7. Gaussian Mixture approximation of the Admissible Region

Figure 5.7 (a), shows the location of the components of the Gaussian Mixture approximation. The design parameters used are $\sigma_\rho = 1000, \sigma_{\dot{\rho}} = 0.95$. Figure 5.7 (b),shows the probability distribution of the Gaussian Mixture. From the plot, it can be seen that the approximation gives a uniform distribution.

5.4.3 Inclusion in the GM-PHD filter

Equation 5.50 gives the GM approximation of the Admissible Region. However, this GM is defined in a 2D space(range,range-rate space) i.e., the means contain $\rho, \dot{\rho}$. In order to use this birth model in the GM-PHD filter the Gaussian components have to be translated into the state space(see equation 4.65). The following steps details the process to do that.

- A GM in the space of the variables $\mathcal{Z} = [\alpha, \dot{\alpha}, \delta, \dot{\delta}, \rho, \dot{\rho}]$ can be obtained as follows:

$$p(\mathcal{Z}) = \sum_{i=1}^{L} \alpha_{\rho,\dot{\rho},i} \mathcal{N}(\mathcal{Z}; m_{\mathcal{Z},i}, P_{\mathcal{Z},i}) \tag{5.51}$$

where

$$m_{\mathcal{Z},i} = \begin{bmatrix} \alpha \\ \delta \\ \rho_i \\ \dot{\alpha} \\ \dot{\delta} \\ \dot{\rho}_i \end{bmatrix}, P_{\mathcal{Z},i} = \begin{bmatrix} \sigma_\alpha^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_\delta^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & P_{\rho_i} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\dot{\alpha}}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\dot{\delta}}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & P_{\dot{\rho}_i} \end{bmatrix} \tag{5.52}$$

$$m_{\rho,\dot{\rho},i} = \begin{bmatrix} \rho_i \\ \dot{\rho}_i \end{bmatrix}, P_{\rho,\dot{\rho},i} = \begin{bmatrix} P_{\rho_i} & 0 \\ 0 & P_{\dot{\rho}_i} \end{bmatrix} \tag{5.53}$$

Note that the terms $\alpha, \dot{\alpha}, \delta, \dot{\delta}$ are obtained from the measurement and its error is $\sigma_\alpha^2, \sigma_\delta^2, \sigma_{\dot{\alpha}}^2, \sigma_{\dot{\delta}}^2$.

- The next step is to translate this to the state space. Since the transformation between $\mathcal{Z}$ and X(state) is non-linear,only an approximate GM in the state space can be obtained by using a linear approximation(for complete proof of the procedure,refer [46]). The birth PHD is then given as:

$$b(\mathbf{x}) = \sum_{i=1}^{L} \alpha_{\rho,\dot{\rho},i} \mathcal{N}(\mathbf{x}; m_i, P_i) \tag{5.54}$$

where m_i, P_i are given as follows.

$$m_i = f(\mathcal{Z}) + \begin{bmatrix} \vec{\mathbf{q}} & \dot{\vec{\mathbf{q}}} \end{bmatrix}^T \tag{5.55}$$

$$f(\mathcal{Z}) = \begin{bmatrix} \rho_i \cos(\delta)\cos(\alpha) \\ \rho_i \cos(\delta)\sin(\alpha) \\ \rho_i \sin(\delta) \\ \dot{\rho}_i \cos(\delta)\cos(\alpha) - \rho\dot{\delta}\sin(\delta)\cos(\alpha) - \rho\dot{\alpha}\cos(\delta)\sin(\alpha) \\ \dot{\rho}_i \cos(\delta)\sin(\alpha) - \rho\dot{\delta}\sin(\delta)\sin(\alpha) + \rho\dot{\alpha}\cos(\delta)\cos(\alpha) \\ \dot{\rho}_i \sin(\delta) + \rho\dot{\delta}\cos(\delta) \end{bmatrix} \tag{5.56}$$

$$P_i = J_{\mathcal{Z},\mathcal{X}} P_{\mathcal{Z},i} J_{\mathcal{Z},\mathcal{X}}^T \tag{5.57}$$

$$J_{\mathcal{Z},\mathcal{X}} = \frac{\partial f(\mathcal{Z})}{\partial \mathcal{Z}} \tag{5.58}$$

The first matrix is obtained from directly from the relations between the variables $\mathcal{Z}$ and X. The state of the observer($\begin{bmatrix} \vec{\mathbf{q}} \\ \dot{\vec{\mathbf{q}}} \end{bmatrix}$) is added since the states of the targets have to be taken in a inertial frame i.e., the Geocentric Equitorial frame. The matrix $J_{\mathcal{Z},\mathcal{X}}$ is known as a Jacobian.

The above steps detail the procedure to obtain the Birth PHD from the GM approximation of an Admissible Region. The next step is to incorporate it into the GM-PHD filter.

Am important point to note that is that the Birth-PHD GM components are all obtained from the measurements. Hence if these are added in the prediction step as shown in equation 4.67 the update step would increase the weights of these components even if these components don't represent the targets. This is because all the components would in theory give the same predicted measurement which is equal to the true measurement. Therefore to avoid this ambiguity, the Birth-PHD GM components are added in the update step which is same as giving new initial components at very time step. Technically, it is then equivalent to assume that there no new objects entering the scene but the new targets(birth) are accounted by adding new GM components(which have information on new objects) at each time step. The GM-PHD filter equations can then be written as:

$$D_{k+1|k}(\mathbf{x}) = D_{S,k+1|k}(\mathbf{x}) \tag{5.59}$$

$$D_{k+1|k+1}(\mathbf{x}) = (1 - p_{D,k+1})D_{k|k-1}(\mathbf{x}) + \sum_{\mathbf{z}\in Z_{k+1}} D_{D,k+1}(\mathbf{x};\mathbf{z}) + b_{k+1|k}(\mathbf{x}) \tag{5.60}$$

A complete definition of the above terms can be found in section 4.3.2. The next chapter provides a method to incorporate the uncertainty in p_D in GM-PHD filter, the research focus of this thesis.

6. UNCERTAINTY IN THE PROBABILITY OF DETECTION

This chapter dives into an important component of the the GM-PHD filter, namely the probability of detection(p_D).As was discussed in the section 4.3.1, p_D is assumed to be constant in the GM-PHD filter. However, that is generally not the case in the space problem.This is because p_D depends on the amount of light reflected by the targets and since the targets are always moving the amount of light reflected changes and this leads to varying p_D. Hence, to model the p_D correctly, the light received by the detector from the target and the signal conversion of this light by the detector is taken into account. Currently, most of the optical measurements (angles and angle rates: $\alpha, \dot{\alpha}, \delta, \dot{\delta}$) are taken with the help of CCD (Charged Couple Device) technology. The following sections briefly describe the signal conversion in the CCD sensor and finally give an analytical expression for p_D. A detailed discussion of the following material can be found in [47] and [24].

6.1 Analytical expression for p_D

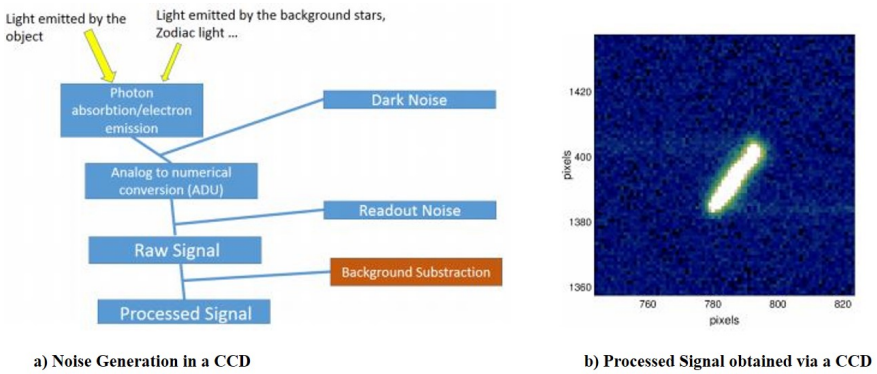


Figure 6.1. Signal Conversion in a CCD sensor(Images taken from [47])

Figure 6.1 illustrates the signal conversion in a CCD sensor and also lists the various sources of noise in the process. The light collected by the CCD sensor falls on semi-conductors present in the sensor which leads to electron emission. However, a *dark noise* is added here because there is a small amount of extra electron emission because of thermal energy of the sensor (since it is never at 0 K). The electrons are then passed over some electronics and signal is amplified to estimate the number of electrons emitted. A *readout noise* is added here in lieu of the electron losses or spurious electron emissions because of flaws in the CCD sensor. The processed signal then looks like Figure 6.1,(b). Note that the light collected by the sensor includes the target signal and the light obtained from the background sources (Example: Stars). Hence, the signal is further processed to eliminate the background sources.
In order to obtain a analytical expression for p_D certain modelling assumptions corresponding to the intensity of the signals are made. They are:

- The number of electrons emitted is assumed to be a Poisson random variable. Then, target signal has a Poisson parameter λ_{obj}, signal from the background sources has a Poisson parameter λ_S and dark noise has a Poisson parameter λ_D.
- The readout noise (S_R) is assumed to be a Gaussian random variable with 0 mean and variance σ_R^2.

The probability of detection (p_D) is then determined by taking the intensity in the brightest pixel of the image and comparing it against a threshold. Therefore,

$$p_D = 1 - P(S_{brightest} < t) \tag{6.1}$$

where $P(S_{brightest} < t)$ is the probability that the intensity of the object signal in the brightest pixel is less than the threshold t. The background signal is estimated by using a average value of signal over the total number of pixels (say n_B). The Background signal for a large value of n_B can then be approximated by a Gaussian random variable with mean $\mu_B = \frac{\lambda_S+\lambda_D}{g}$ and variance $\sigma_B^2 = \frac{\lambda_S+\lambda_D}{gn_B}$ where g is the gain

used to amplify the electron signal. A detailed procedure can be found in [47].Then, the probability of detection can be given by:

$$p_D = 1 - \frac{1}{2}\sum_{n=-\infty}^{\infty} \frac{\Gamma(n+1, \lambda_{obj}+\lambda_S+\lambda_D)}{n!}\left(erf\left(\frac{n+1-\mu_\epsilon}{\sqrt{2g(\sigma_B^2+\sigma_R^2)}}\right) - erf\left(\frac{n-\mu_\epsilon}{\sqrt{2g(\sigma_B^2+\sigma_R^2)}}\right)\right) \tag{6.2}$$

$$\mu_\epsilon = tg + g/2 + \lambda_S + \lambda_D \tag{6.3}$$

$$t = 3 \times N(\text{ threshold, usually equal to 3-5 N}) \tag{6.4}$$

$$N = \sqrt{\frac{\lambda_{obj} + (1+\frac{1}{n})(\lambda_S + \frac{g^2-1}{12})}{g^2} + (1+\frac{1}{n})\sigma_R^2} \tag{6.5}$$

where erf is the error function. A complete derivation can be found in [47]. **Note** : The incomplete gamma function present in the above equation is numerically equal to the cdf of a poisson random variable with parameter $\lambda = \lambda_{obj} + \lambda_S + \lambda_D$ as follows:

$$\frac{\Gamma(n+1, \lambda_{obj}+\lambda_S+\lambda_D)}{n!} = \sum_{k=0}^{n} \frac{\exp(-\lambda)\lambda^k}{k!} = \text{cdf of poisson distribution} = poisson.cdf(n, \lambda)$$

The next section provides a brief analysis of the object signal.

6.2 Object Irradiation

As discussed in the previous section, the amount of light reflected by a target depends on the relative geometry between the object, the observer and the illumination source (Sun) and the reflecting surface of the object. The object irradiation at the location of observer(in W/m^2) is then given by (refer [24]) :

$$I_{obj} = \int I_{Sun}(\lambda)\frac{A}{x_{topo}^2}\bar{\Psi}(\lambda)d\lambda \equiv \frac{I_0 A \Psi}{r_{topo}^2} \tag{6.6}$$

where I_{Sun} is the Sun's intensity at the object's location for a given wavelength λ. Generally, $I_{Sun}\Delta\lambda$ is approximated by the solar constant $I_0 = 1367.7W/m^2$ where λ is taken as the mean wavelength. A is the total area of the reflecting surface, r_{topo} is the topocentric distance (distance between the object and observer) and Ψ is the known as the phase function. The phase function depends on the object properties, specifically

the reflecting surface i.e., encapsulates the information of the shape and reflection properties of the object. For a spherical object, Ψ is given as follows (refer [42] for complete proof):

$$\Psi_{sphere} = \frac{2C_d}{3\pi}(\sin(\alpha) + (\pi - \alpha)\cos(\alpha)) \tag{6.7}$$

where C_d is the Lambertian Reflection co-efficient and α is the angle between the relative position vector between Sun and the object and the relative position vector between the object and the observer.C_d is generally assumed to be a constant and value is taken at the mean wavelength. The objects for the purpose of this thesis are assumed to spherical because of the analytical expression available for spheres.

6.2.1 CCD sensor response

When this irradiation passes through the optics in a CCD sensor, the expected value for the signal function can be given as follows (refer [42]):

$$E(I_{obj}) = \int (D-d)\frac{\lambda}{hc} I_{\lambda obj}(\lambda)\exp(-\tau(\lambda)R(\zeta))d\lambda \tag{6.8}$$

where D is the area of the aperture, d is the area of the obstruction of the aperture(secondary mirror), c is the speed of light, h is the Planck's constant, $I_{\lambda obj}(\lambda)$ is the object irradiation for a wavelength λ,τ is the atmospheric extinction coefficient and R is the atmospheric function (ζ is the elevation angle). The atmospheric model used in this thesis is $R = \frac{1}{\cos(\zeta)}$. The equation is simplified by evaluating $E(I_{obj})$ at the mean wavelength $\bar{\lambda}$ and is given by:

$$E(I_{obj}) = (D-d)\frac{\bar{\lambda}}{hc} I_{obj}(\bar{\lambda})\exp(-\tau(\bar{\lambda})R(\zeta)) \tag{6.9}$$

The number of electrons ejected i.e., λ_{obj} is then given by:

$$\lambda_{obj} = (D-d)\frac{\bar{\lambda}}{hc} I_{obj}(\bar{\lambda})\exp(-\tau(\bar{\lambda})R(\zeta))Q(\bar{\lambda})\Delta t g \tag{6.10}$$

where Q is the quantum efficiency and g is the gain.λ_{obj} can be rewritten as:

$$\lambda_{obj} = k_{obj}A \tag{6.11}$$

where k_{obj} is a number which is independent of area and depends on the location of the object, location of the sun, reflective properties of the object and the properties of the sensor used to obtain the CCD image.

6.3 State Dependent p_D

As discussed in the previous section, constant p_D assumption is not valid in the case of the space problem. From equation 6.10, it can be seen that λ_{obj} depends on the state(position) of the object which implies that p_D is state dependent. Hence, in order to get a tractable form for the GM-PHD filter equations, [48] suggests an alternative in which the the following assumption is made:

- p_D is approximated as a constant around the mean of the considered Gaussian component. This assumption is completely valid when when the probability of detection(p_D) varies slowly as compared to the Gaussian components of the PHD in the update step($D_{k+1|k+1}(\mathbf{x})$).

Consider the update step for a PHD filter is given by (equation 4.58):

$$D_{k+1|k+1}(\mathbf{x}) = \left(1 - p_D(\mathbf{x}) + \sum_{\mathbf{z}\in Z_{k+1}} \frac{p_D(\mathbf{x}) f_{k+1}(\mathbf{z}|\mathbf{x})(\mathbf{x})}{\lambda c(\mathbf{z}) + \int p_D(\xi) D_{k+1|k}(\xi) f_{k+1}(\mathbf{z}|\xi)}\right) D_{k+1|k}(\mathbf{x}) \tag{6.12}$$

Then under the assumptions of the GM-PHD filter and the additional assumption made above, the new GM-PHD filter update equation becomes (refer [48] for complete proof):

$$D_{k+1|k+1}(\mathbf{x}) = [1 - p_D] D_{k+1|k}(\mathbf{x}) + \sum_{\mathbf{z}\in Z_{k+1}} D_{D,k+1}(\mathbf{x};\mathbf{z}) \tag{6.13}$$

where,

$$[1 - p_D] D_{k+1|k}(\mathbf{x}) = \sum_{i=1}^{J_{k+1|k}} [1 - p_{D,k+1}(m^{(i)}_{k+1|k})] w^{(i)}_{k+1|k} \mathcal{N}(x; m^{(i)}_{k+1|k}, P^{(i)}_{k+1|k}) \tag{6.14}$$

$$D_{D,k+1}(\mathbf{x};\mathbf{z}) = \frac{p_{D,k+1}(m^{(i)}_{k+1|k}) w^{(i)}_{k+1|k} q^{(i)}_{k+1}(z)}{\lambda c(\mathbf{z}) + p_{D,k+1}(m^{(i)}_{k+1|k}) \sum_{l=1}^{J_{k+1|k}} w^{(l)}_{k+1|k} q^{(l)}_{k+1}(\mathbf{z})} \tag{6.15}$$

The following points briefly explain the derivation of the above equations.

- In the first term $((1 - p_D(\mathbf{x}))D_{k+1|k}(\mathbf{x}))$ in equation 6.12, $p_D(\mathbf{x})$ converts to $p_D(m)$ because, according to the assumption, p_D varies slowly as compared to the variation of the Gaussian term i.e., for locations farther from the mean of the Gaussian the term $p_D(x)v_{k|k-1}(x)$ is approximately zero because $v_{k|k-1}(x)$ is approximately zero and p_D is a value between 0 and 1.

- In the second term of equation 6.12, the numerator is obtained in a similar way as described in the previous point. The denominator is obtained by using the constant p_D assumption in the following way :

$$\int p_D(x)\mathcal{N}(x; m, P)dx \approx p_D(m) \int \mathcal{N}(x; m, P)dx \tag{6.16}$$

- As discussed previously the $p_{D,k+1}(m^{(i)}_{k+1|k})$ is then obtained by using the CCD equation (equation 6.2). But if the FOV is restricted as described by section 5.3.1, then $p_{D,k+1}(m^{(i)}_{k+1|k})$ is taken to be zero if the estimated position (obtained from $m^{(i)}_{k+1|k}$) is not in the FOV.

6.4 Uncertainty in the Probability of Detection

This section provides a method to incorporate the uncertainty in p_D in a GM-PHD filter. From equations 6.2 and 6.10 it can be seen that p_D depends on the area of the reflecting surface. For the objects moving in space, especially the debris, since they are in an uncontrolled motion, their orientation is difficult to evaluate and hence it makes it difficult to estimate the reflecting area. This section introduces a method to integrate the uncertainty in p_D in the GM-PHD filter equations when the area of the targets can be assumed to have a Gaussian distribution.

Let Area be a Gaussian random variable with a mean μ_A and variance σ_A^2. Since p_D has a non-linear relationship with area, the probability distribution for p_D is not

exactly a Gaussian. Hence,the probability distribution of p_D is linearly approximated by a Gaussian distribution with parameters as follows:

$$\mu_{p_D} = p_D(\mu_A) \text{ ,i.e., use Area} = \mu_A \text{ to evaluate } \lambda_{obj} \tag{6.17}$$

$$\sigma^2_{p_D} = \left(\left. \frac{dp_D}{dA} \right|_{A=\mu_A} \sigma_A \right)^2 \tag{6.18}$$

$\left. \frac{dp_D}{dA} \right|_{A=\mu_A}$ is known as the Jacobian and is calculated as follows:

- In order to simplify the process of finding the derivative, p_D is divided into two parts as follows:

$$p_D = 1 - \frac{1}{2} \sum_{n=-\infty}^{\infty} P_1.P_2$$

$$P_1 = poisson.cdf(n, \lambda) \text{ where } \lambda = \lambda_{obj} + \lambda_S + \lambda_D$$

$$P_2 = erf\left(\frac{n+1-\mu_\epsilon}{\sqrt{2g(\sigma_B^2 + \sigma_R^2)}} \right) - erf\left(\frac{n-\mu_\epsilon}{\sqrt{2g(\sigma_B^2 + \sigma_R^2)}} \right)$$

  Then, the derivative can be written as:

$$\frac{dp_D}{dA} = -0.5 \sum_{n=-\infty}^{\infty} \left(\frac{dP_1}{dA} P_2 + P_1 \frac{dP_2}{dA} \right)$$

- The term $\frac{dP_1}{dA}$ is obtained by using the chain rule of derivation and is as follows:

$$\frac{dP_1}{dA} = k_{obj}(poisson.cdf(n-1, \lambda) - poisson.cdf(n, \lambda)) = \frac{-k_{obj} \exp(\lambda) \lambda^n}{n!}$$

- The differentiation of P_2 is a bit more involved since the area dependence is through the threshold(t) term and also since it contains an error function. Hence, the Leibnitz rule needs to be applied which is given as follows:

$$\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t)dt \right) = f(x, b(x)) \frac{db(x)}{dx} - f(x, a(x)) \frac{da(x)}{dx} + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)dt$$

- Before proceeding further, define:

$$c = \sqrt{2g(\sigma_B^2 + \sigma_R^2)}$$

$$z_1 = \frac{n+1-\mu_\epsilon}{c}$$

$$z_2 = \frac{n-\mu_\epsilon}{c}$$

Then the term P_2 is:

$$P_2 = erf(z_1) - erf(z_2)$$

- Applying the Leibnitz rule on the terms of P_2 yields:

$$\frac{d(erf(z_1))}{dA} = \frac{2}{\sqrt{\pi}} \exp(-z_1^2) \frac{dz_1}{dA} = \frac{2}{\sqrt{\pi}} \exp(-z_1^2) \frac{-g}{c} \frac{dt}{dA}$$
$$\frac{d(erf(z_2))}{dA} = \frac{2}{\sqrt{\pi}} \exp(-z_2^2) \frac{dz_2}{dA} = \frac{2}{\sqrt{\pi}} \exp(-z_2^2) \frac{-g}{c} \frac{dt}{dA}$$

where:

$$\frac{dt}{dA} = \frac{1.5 k_{obj}}{g^2} \left(\frac{\lambda_{obj} + (1 + \frac{1}{n})(\lambda_S + \frac{g^2 - 1}{12})}{g^2} + (1 + \frac{1}{n})\sigma_R^2 \right)^{-0.5}$$

- The Jacobian $\frac{dp_D}{dA}$ can be found by combining all the equations above.

The following figure shows the variation of p_D for an object in a Geostationary orbit as it moves along it's orbit. The object is assumed to be spherical and the Area is assumed to be equal to 1.

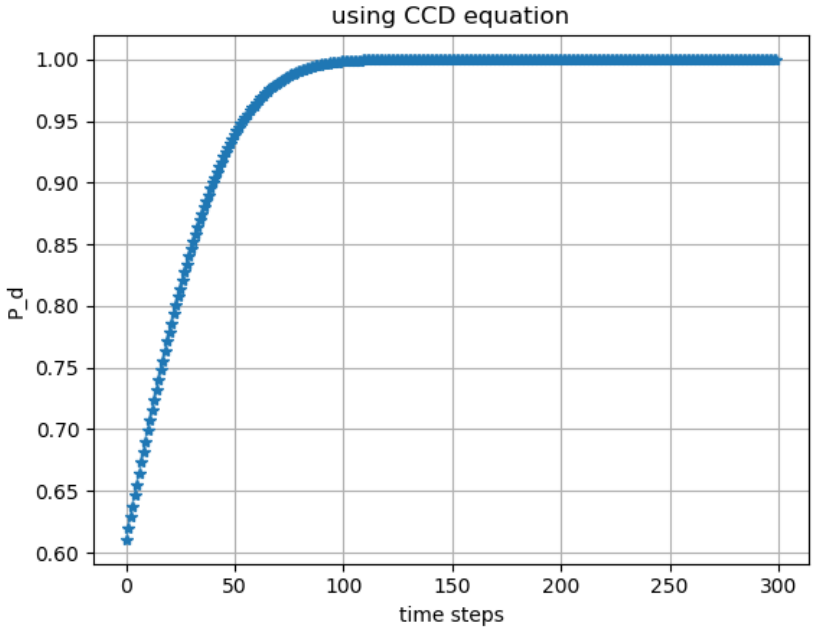


Figure 6.2. p_D vs time for a Geostationary object

In order to verify the GM approximation of p_D, a Monte-Carlo analysis is done. The process is as follows:

- The probability distribution of the area is assumed to be Gaussian with $\mu, \sigma = 1, 0.3$.

- To test how good the Gaussian distribution fits the p_D distribution 1000 samples of Area are drawn and p_D is calculated for each of the sample. A histogram is plotted using 10 bins where the bin intervals are [0-0.1,0.1-0.2,...,0.9-1.0].

- Then the probability of p_D at the mid point of bin intervals [0.05,0.15,...,0.85,0.95] is calculated by taking the ratio of the number of samples in each bin to the total number of samples (1000).

- The cdf (cumulative distribution function, sum of the probabilities over a particular interval) of p_D using the Gaussian model is also calculated in each of the bin intervals and this value is assigned as the probability for the mid-point of bin intervals. Also, since $0 \leq p_D \leq 1$ the pdf at $p_D = 0, 1$ are modified in the following way.

$$p(p_D = 0.05) = p(p_D = 0.05) + p(p_D < 0) \tag{6.19}$$

$$p(p_D = 0.95) = p(p_D = 0.95) + p(p_D > 1) \tag{6.20}$$

- The following figure (figure 6.3) shows the comparison of probability of p_D at the bin mid-points obtained via Monte-Carlo analysis and the Gaussian approximation. It can be seen from the plots that the Gaussian approximation is a very good one when p_D(for Area = 1) is high i.e., at higher time steps (see figure 6.2).

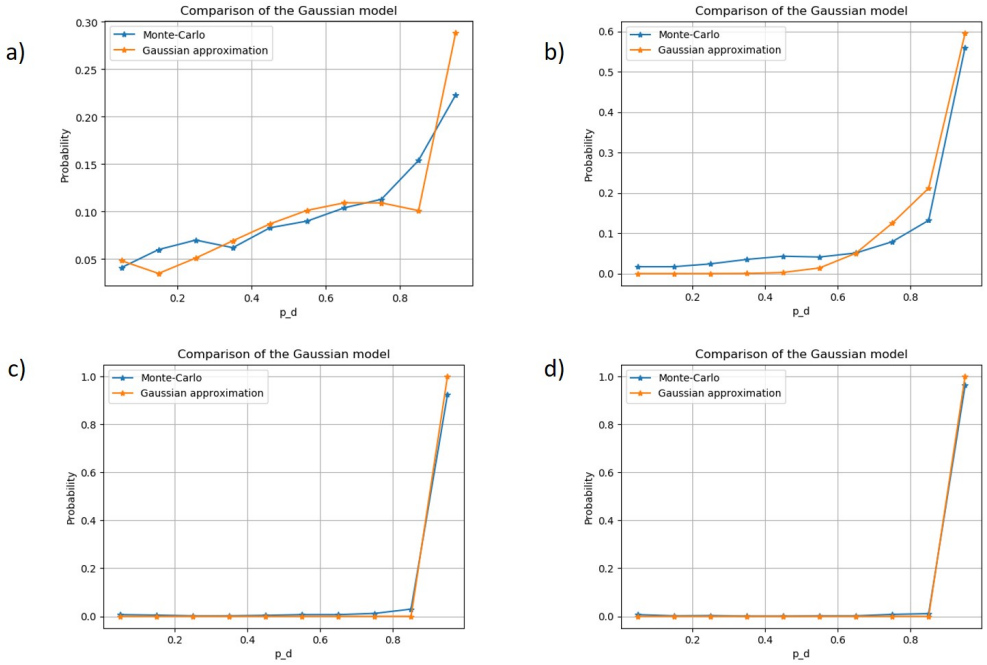


Figure 6.3. a) t = 10, b) t = 50, c) t = 150, d) t = 200

A better approximation is possible by splitting the Gaussian distribution of Area into multiple Gaussians and then use the approximation strategy on the split Gaussians. The following section gives a brief overview of that process.

6.4.1 Splitting a Gaussian distribution

The following points systematically describe the process to split a Gaussian distribution into a Gaussian Mixture.

- Let $p(x)$ be the Gaussian distribution that is sought to be split. Without loss of generality, it can be assumed that it is a standard normal distribution.

$$p(x) = \mathcal{N}(x; 0, 1)$$

- The objective is to approximate $p(x)$ by a Gaussian mixture $\tilde{p}(x)$ which is expressed as:

$$\tilde{p}(x) = \sum_{i=1}^{n} \alpha_i \mathcal{N}(x; \mu_i, \tilde{\sigma}^2)$$

where n is the number of components of the Gaussian mixture.

- In order to reduce the number of variables to be evaluated it is assumed that all the components have the same variance i.e., the Gaussian mixture is homoscedastic.

- The parameters of the GMM are solved for by a posing a minimisation/optimisation problem i.e., L_2 norm (shown below) between $p(x)$ and $\tilde{p}(x)$ is minimised.

$$D = \frac{1}{2} \int_{-\infty}^{\infty} (p(x) - \tilde{p}(x))^2 dx$$

The L_2 norm is chosen because the divergence (J) between $p(x)$ and $p(\tilde{x})$ can be found in a closed form.

- Since the objective is to find GMM components with smaller variances a penalty term λ is added which weighs the importance of minimising the divergence with the minimising of variance ($\tilde{\sigma}^2$).The optimisation problem is then stated as :

$$\min J + \lambda \tilde{\sigma}^2 \text{ subject to } \sum_{i=1}^{n} \alpha_i = 1$$

- By solving the above optimisation problem, splitting libraries can be formed by varying n, λ. The following figure(figure 6.4) show the pre-computed results from [49], [50] for n=3,4,5

- The following table shows the library of solution for n=4 which was computed by authors of [50]

Table 1 Three-component splitting library with $\beta = 2$ and $\lambda = 0.001$

i	$\tilde{\alpha}_i$	$\tilde{m}_i$	$\tilde{\sigma}$
1	0.2252246249	−1.0575154615	0.6715662887
2	0.5495507502	0	0.6715662887
3	0.2252246249	1.0575154615	0.6715662887

Table 2 Five-component splitting library with $\beta = 2$ and $\lambda = 0.0025$

i	$\tilde{\alpha}_i$	$\tilde{m}_i$	$\tilde{\sigma}$
1	0.0763216491	−1.6899729111	0.4422555386
2	0.2474417860	−0.8009283834	0.4422555386
3	0.3524731300	0	0.4422555386
4	0.2474417860	0.8009283834	0.4422555386
5	0.0763216491	1.6899729111	0.4422555386

Figure 6.4. Splitting Library(Source : pg 1051 of [49])

SPLITTING LIBRARY

i	$\tilde{\omega}_i$	$\tilde{\mu}_i$	$\tilde{\sigma}_i$
1	0.12738084098	-1.4131205233	0.51751260421
2	0.37261915901	-0.44973059608	0.51751260421
3	0.37261915901	0.44973059608	0.51751260421
4	0.12738084098	1.4131205233	0.51751260421

Figure 6.5. Splitting Library(Source: pg 183 of [50])

- It is important to note that the solution libraries that are obtained are for a standard normal distribution. For a normal distribution with $\mu \neq 0, \sigma \neq 1$, the parameters of each component can be evaluated as follows:

$$\mu_i = \mu + \tilde{\sigma}\tilde{\mu}_i$$

$$\sigma_i = \tilde{\sigma}_i\sigma$$

$$\alpha_i = \tilde{\alpha}_i$$

- The following figures compare the results after splitting the Gaussian Mixture into 3, 4 and 5 components.

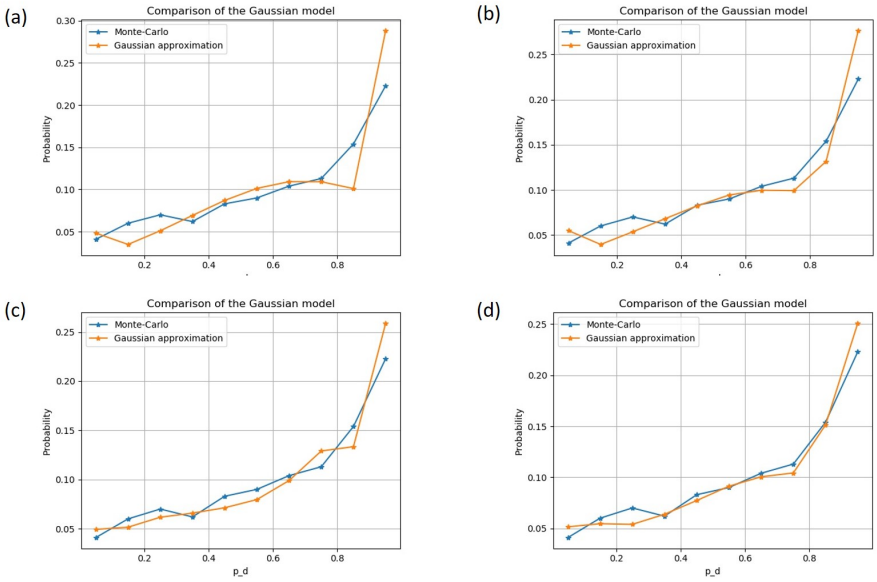


Figure 6.6. (a) n = 1,(b) n = 3,(c) n = 4,(d) n = 5 at t=10

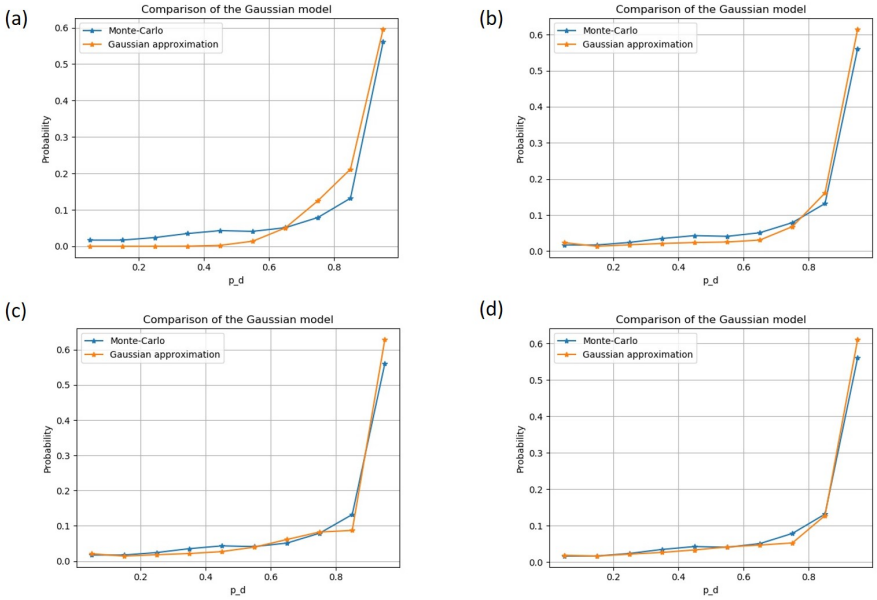


Figure 6.7. (a) n = 1,(b) n = 3,(c) n = 4,(d) n = 5 at t=50

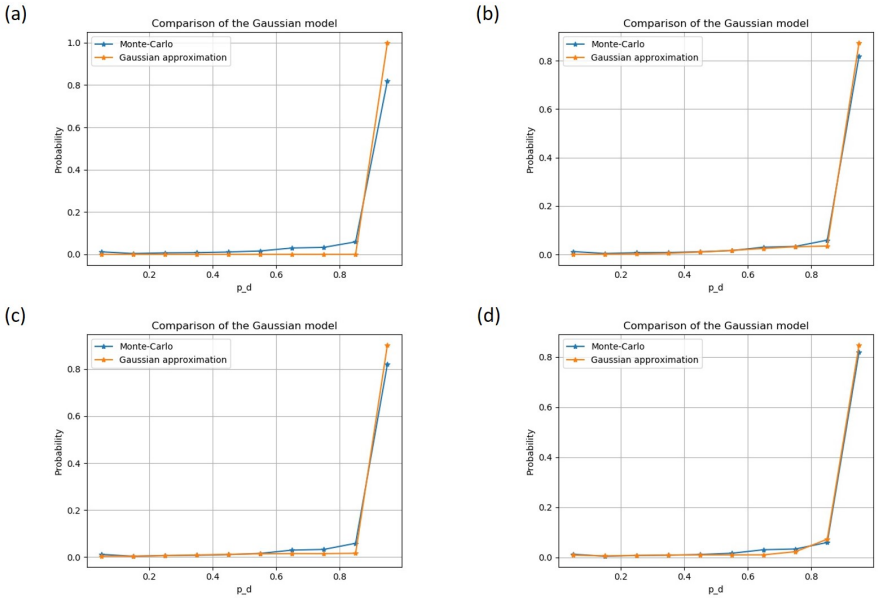


Figure 6.8. (a) n = 1,(b) n = 3,(c) n = 4,(d) n = 5 at t=100

From the above results it can be seen that the Gaussian approximation gets better as more components are used.

For the implementation in the GM-PHD filter, $p_D(m^{(i)}_{k+1|k})$ in equations 6.14 and 6.15 is taken as the weighted average of the means of the Gaussian components i.e., the splitted Gaussian components of Area are first transformed to p_D space by the method described in section 6.4 and the p_D is assumed to then be equal to the weighted average of the means of the components. More precisely, if the probability distribution of p_D for a particular state $m^{(i)}_{k+1|k}$ is given by $\sum_{i=1}^{n} \alpha_i \mathcal{N}(p_D; \mu_i, \sigma_i^2)$ then $p_D(m^{(i)}_{k+1|k}) = \sum_{i=1}^{n} \alpha_i \mu_i$. The results of this method will be investigated in detail in the next chapter. The following section introduces the OSPA metric which will be used to validate the p_D uncertainty model.

6.5 OSPA metric

In the GM-PHD filter two main quantities are estimated at each time step, i.e., the cardinality and the state of the targets. Hence, the performance of the filter can be

assessed accurately by considering the error in cardinality and the error in estimation of state. In [51], many such metrics are looked at and the inconsistencies in each metric are outlined and a metric (OSPA) which addresses the drawbacks of the other multi-object performance evaluation metrics is proposed. The following expression gives the OSPA metric(detailed discussion can be found in [51]) :

$$\bar{d}_p^{(c)}(X,Y) := \left(\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}(x_i, y_{\pi(i)})^p + c^p(n-m) \right) \right)^{1/p}$$

In the above equation the first term($\min_{\pi \in \Pi_n} \sum_{i=1}^{m} d^{(c)}(x_i, y_{\pi(i)})$) is obtained by optimal sub-pattern assignment which can be described as follows:

- Consider two sets X (truth) and Y (estimate) with cardinalities n and m respectively (without loss of generality,assume m $\leq$ n). In this thesis, these sets will contain the positions of the targets.

- The distances between the elements of each set can then be calculated and matrix (size mxn) containing the distances can be obtained. Also, the parameter c is defined to remove any outliers i.e, if any distance is greater than c the calculated distance is replaced by c. The optimal sub-pattern assignment would then be to assign the closest n out of m estimates to correspond to the n truths. The first term is then the sum of these n distances.

- The second term calculates the error in cardinality by multiplying the difference in number of estimated and true targets by the parameter c.

The choice of the parameters c, p is problem dependent. For the purposes of this thesis, $c = 10000, p = 1$.

7. RESULTS

The necessary background required for the implementation of the GM-PHD filter in a SSA scenario has been discussed until now. It makes sense to use the GM-PHD filter in a SSA scenario, because firstly, it provides a closed-form solution and is also a computationally tractable and a less expensive algorithm. However, in its original form, as developed by Vo et.al [1], has some limitations in the context of SSA. p_D is assumed to be state independent at any time i.e., it is assumed to be a constant. As discussed in chapter 1 and chapter 6, in the space environment, p_D is state dependent and is also uncertain. It is state dependent because p_D depends on the relative position of the object with respect to the observer and the Sun and the uncertainty is introduced because of the degradation of the surface materials and also the uncertain estimation of the object's reflecting area. Hence, a state dependent p_D model, relevant to SSA, was introduced in chapter 6. Using this model, variations of GM-PHD filter which accommodate a state dependent p_D (section 6.3) and a uncertainty in p_D (section 6.4) have been developed and their performance is tested in this section.

The models and assumptions used in the GM-PHD filter were described in chapter 5. A single ground based sensor is used as the source of measurements. Only optical measurements are considered. The results including and excluding the clutter measurements are shown. The target measurement model and the target motion model have been described in chapter 5. For the track initialization, i.e., the birth model for the GM-PHD filter (see section 4.3), Admissible Regions are used. A complete description of this model is given in section 5.4.2. The OSPA metric used to validate the p_D uncertainty model is described in section 6.5.

This chapter is divided into four sections. Section 7.1 shows the application of the basic GM-PHD filter on a SSA problem assuming that the probability of detection is constant. Section 7.1 is further divided into two sections. In section 7.1.1, it is

assumed that there is no FOV restriction i.e., the observer on the Earth can observe any object at any time. Though the constant p_D assumption and the unrestricted FOV assumptions are unrealistic, section 7.1.1 is included to show the performance of the GM-PHD filter with respect to the availability of measurements. In section 7.1.2, the FOV restriction is included, which is more realistic and the filter performance is assessed.

Section 7.2 presents the filter performance by changing the p_D model, i.e., the probability of detection (p_D) is assumed to be state dependent and uses a analytical expression for p_D taking various factors into account (described in detail in chapter 6).

Sections 7.1 and 7.2 assumes that there is model mismatch in the p_D i.e., the filter p_D model is assumed to be the truth. This is not necessarily true always since it not possible to model the reality perfectly. This section assesses the filter's performance in case of a model mismatch and signifies the need of introducing an uncertainty in the GM-PHD filter.

Section 7.4 validates a p_D model (introduced in section 6.4) which takes an uncertainty in p_D introduced by an uncertainty in Area into account and incorporates it into the GM-PHD filter.

7.1 Constant p_D GM-PHD filter

This section demonstrates the application of the GM-PHD filter assuming that probability of detection is constant, there is no FOV restriction (targets generate measurements always) and there is no model mismatch i.e., the true p_D and the p_D that the filter uses are the same. This is a very simple case and PHD filter is known to deal best with high p_D and the unrestricted field of view allowed for continuous measurements. The performance of the GM-PHD filter is studied by taking a low (0.5) and a high (0.9) value of p_D.

7.1.1 Unrestricted Field of View

The simulation is run for five objects for a total of 300 time steps (each time step is 60 seconds) where measurements are taken at every time step and the time of "birth" and "death" of the targets are shown in figure 7.1.1. The orbital elements of the targets are shown in figure 7.1.1. It is assumed that the measurements are unconstrained i.e., there is no Field of View restriction and the observer's i.e., the ground based sensor's Geocentric Equitorial coordinates at $t = 0$ are $\alpha = 0°, \delta = 45°$, which is a location in France. Measurements are taken every 60 seconds and consist of topocentric angles and angle-rates $(\alpha, \delta, \dot{\alpha}, \dot{\delta})$. It is assumed that the clutter Poisson parameter is 4 (λ). To show the effect of clutter, the results assuming measurements with no clutter are also presented in this section. Probability of survival is constant and is assumed to be equal to 0.99. It is assumed that the probability of detection(p_D) is constant and there is no model mismatch i.e., p_D for the truth and p_D used in the filter are the same. For the birth model, the design parameters are assumed as $\sigma_\rho = 2750, \sigma_{\dot{\rho}} = 1.2$ (see section 5.4.2).

Targets	Birth Time Step	Death
Object 1	0	-
Object 2	0	45
Object 3	0	220
Object 4	20	-
Object 5	70	-

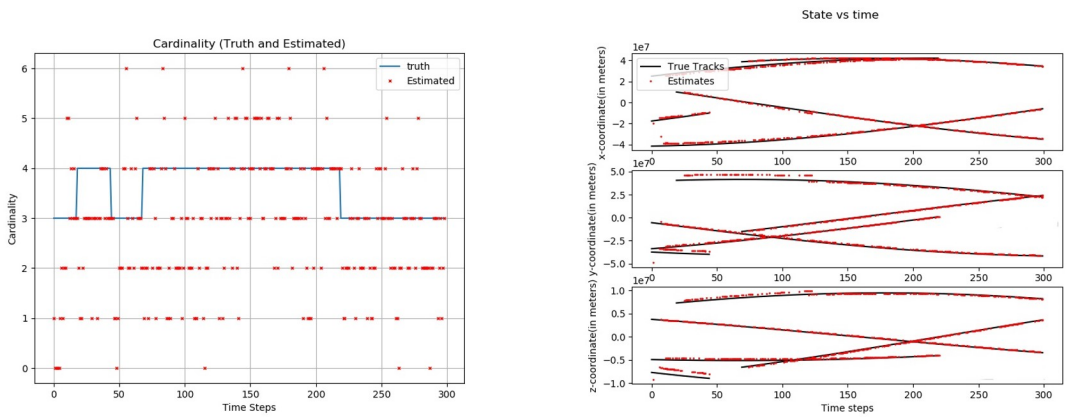
Figure 7.1. Target birth and death time steps

Targets	Semi-major axis (km)	eccentricity	Inclination (radians)	Longitude of ascending node (radians)	Argument of perigee (radians)	True Anomaly (radians)
Object 1	42165.848	0.00039	0.14138	0.81459	0.67919	1.77549
Object 2	42409.159	0.00061	0.22575	0.46481	0.95157	6.20904
Object 3	42168.386	0.00063	0.25615	0.26807	3.94924	1.67642
Object 4	42163.991	0.00023	0.25443	0.34034	1.81697	2.13028
Object 5	42474.792	0.00034	0.12035	0.93499	0.54867	3.86759

Figure 7.2. Orbital Elements of the targets

p_D=0.5, no clutter

Figure 7.3 shows the results when $p_D = 0.5$ and there is no clutter. Since, it is a unrestricted FOV, it means that at any given time, there is 50% chance that a target generates a measurement. From the state estimate plot (part (b)) it can be seen that 3 objects start getting tracked initially. This is because only object 1, 2 and 3 are in the scene initially. Objects 4 and 5 start getting estimated once they are "born". It can be seen that it takes some time for the objects to be estimated, since sufficient measurements have to be obtained first for the objects to be selected as an estimate (see section 4.3.2) i.e., the weight has to be greater than 0.5. Since there are lesser target generated measurements because of low p_D, the number of objects are underestimated routinely (see part (a)). Also, it can be seen that sometimes the number of objects are overestimated too, this is because the first term in the update term which has 1-p_D term in it (see equation 4.95), is counted as an estimate (if the weight is around 1) and if there is a measurement from that particular target (50% chance), the second term generates another estimate at the same location i.e., two objects are estimated at same state instead of one. The cardinality estimates range from 0 to 6 and on an average, the estimates are incorrect by 1. The state estimates are very accurate after a certain period of time. This can be seen in part(b), at the end of most of the tracks the red dots (estimates) directly overlap the true tracks (black lines).

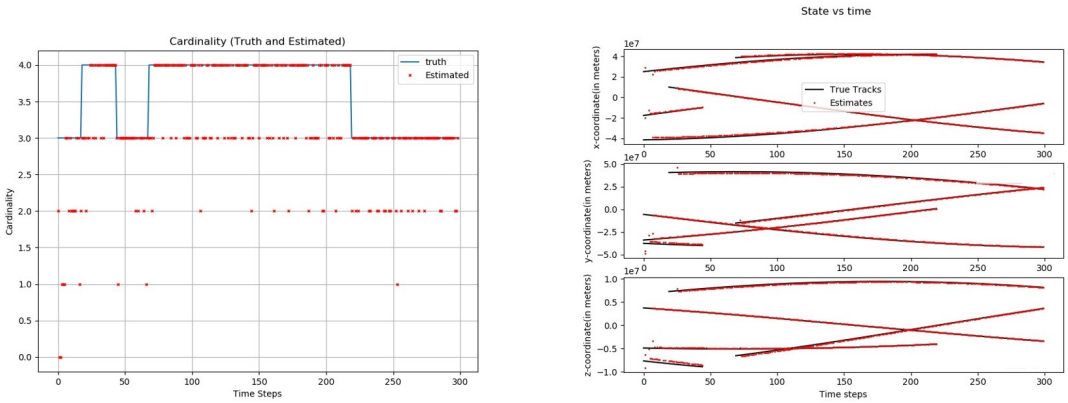


a) Cardinality Plot b) State Estimates

Figure 7.3. For object ground-based tracking scenario, p_D=0.5, no clutter with object birth: a) Cardinality as a function of time, blue the truth, in red the estimates; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z

$\mathbf{p_D}$=0.9, no clutter

Figure 7.4 shows the results when $p_D = 0.9$. It can be seen that the estimates are very close to the truth in this case. This is because of the continuous measurements being available to the filter because of high probability of detection. From part (a) it can be seen that the number of estimates are never overestimated since the first term of equation 4.95, doesn't result in an estimate because $1 - p_D$ is much less and the second term results in an estimate only when there is measurement available (which in this case it is most of the time). The objects are underestimated only when there are no measurements.



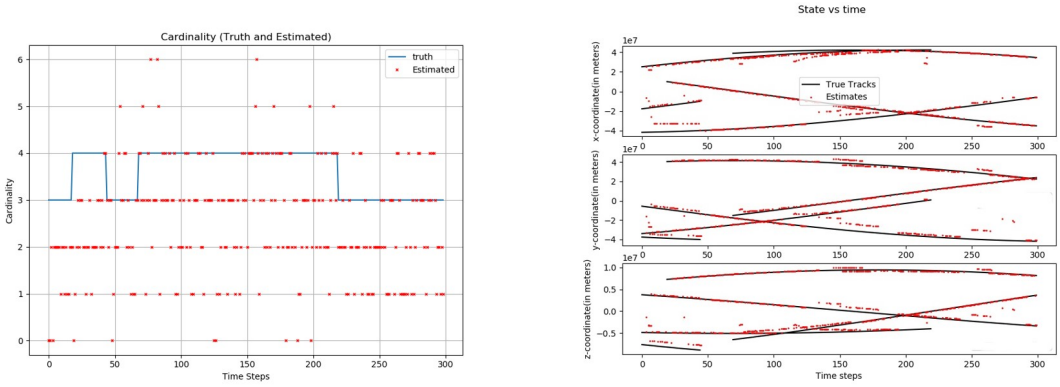
a) Cardinality Plot

b) State Estimates

Figure 7.4. For object ground-based tracking scenario, p_D=0.9, no clutter with object birth: a) Cardinality as a function of time, blue the truth, in red the estimates; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z

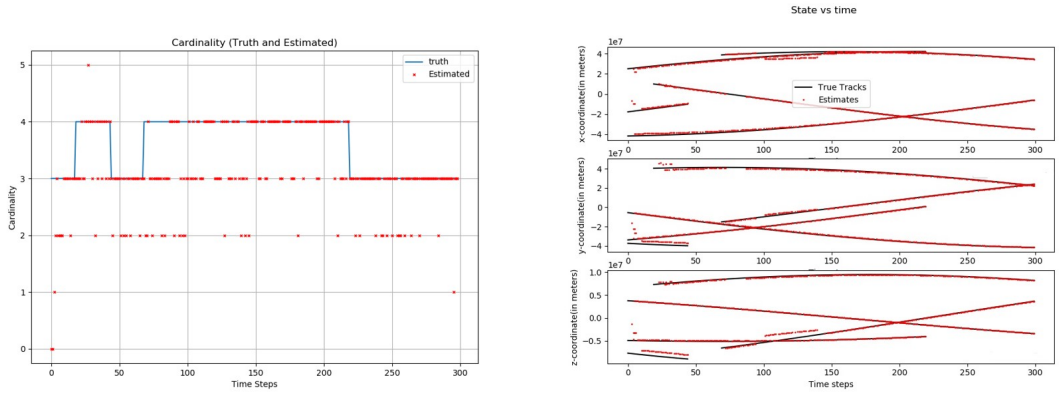
$\mathbf{p_D}$=0.5,0.9, with clutter

The following figures (Figure 7.5 and Figure 7.6) show the results when clutter is present. Since the Poisson paramter is 4, the average measurements rise by 4. Hence, the filter has the additional task of removing the uncertainty in measurements too. Clearly, the filter performance is better when there are no clutter measurements. Consider the case when $p_D = 0.9$. From part (a) of figures 7.4 and 7.9, it can be seen that when clutter is included, there is a significant delay in initializing the track of object 5. Also, the estimate of z coordinates have a higher deviation from the truth when clutter is included.



a) Cardinality Plot b) State Estimates

Figure 7.5. For object ground-based tracking scenario, p_D=0.5, with clutter and object birth: a) Cardinality as a function of time, blue the truth, in red the estimates; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z



a) Cardinality Plot b) State Estimates

Figure 7.6. For object ground-based tracking scenario, p_D=0.9, with clutter and object birth: a) Cardinality as a function of time, blue the truth, in red the estimates; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z

It can be seen that the GM-PHD filter performs efficiently when p_D is high. However, low p_D values are not uncommon in SSA observations. Low p_D can occur even for large objects too if it is not reflecting enough light because of the material's reflecting property. Smaller apertures also lead to lower p_D.

An unrestricted FOV is unrealistic since the FOV is always restricted for a ground-based observer taking optical measurements. For example, an object, can go below the local horizon and there is no way of observing it from that location at that time. To make the simulations more realistic, a constrained FOV is considered and this assumption will be continued throughout the rest of this chapter.

7.1.2 Restricted Field of View

This section presents the results when the FOV is restricted as described in section 5.3.1. This simulation is run for 300 time steps with a step size of 1 minute. The objects time in FOV is shown in figure 7.1.2. The orbit elements are shown in figure 7.1.2. The observer is assumed to be on the surface of the Earth at a $-91.44°$ longitude and $-7.15°$ latitude and measurements are taken at every time step and consist of topocentric angles and angle-rates $(\alpha, \delta, \dot{\alpha}, \dot{\delta})$. The FOV is assumed to be $4° \times 4°$. It is assumed that the clutter poisson parameter is 4. Probability of survival is constant and is assumed to be equal to 0.99. It is assumed that the probability of detection (p_D) is constant and there is no model mismatch i.e., p_D for the truth and p_D used in the filter are the same. For the birth model, the design parameters are assumed as $\sigma_\rho = 2750, \sigma_{\dot{\rho}} = 1.2$ (see section 5.4.2).

Targets	Time steps in FOV
Object 1	0 – 177 , 185-196
Object 2	80-147

Figure 7.7. Objects periods in FOV

Targets	Semi-major axis (km)	eccentricity	Inclination (radians)	Longitude of ascending node (radians)	Argument of perigee (radians)	True Anomaly (radians)
Object 1	42164.455	0.00027	0.14278	0.80659	4.50221	0.18799
Object 2	42675.265	0.00178	0.22682	0.24389	4.53186	1.01972

Figure 7.8. Orbital Elements of targets

Figure 7.9 shows the results for the case when p_D is high, i.e., $p_D = 0.9$. Form part (a), it can be seen that initially since only object one is in the FOV, at most one target generated measurement is available. This is reflected in the estimates too since only one object is being estimated until the 80th time step. However, at the 80th time step, object 2 enters the FOV and the number of measurements indicate this. Hence, because of high p_D, measurements are readily available and the filter is able to estimate it efficiently. Also, it can be seen that the filter briefly overshoots (for 2 time steps) the number of objects when object 2 enters the FOV. Around the 147th time step, since object 2 exits the FOV, there are no more target generated measurements from it and hence only object 1 is tracked until it exits the FOV. It can be seen from the plot (a) that the estimates follow the number of measurements 90% of the time i.e., if there are continuous measurements, the filter works efficiently. From the state estimate plot, it can be seen that the estimates are not good initially, but as observations of object 1 increase the filter corrects itself. Also, it can be seen that the x coordinate estimate has the highest deviation from the truth. This is introduced because of the uncertainty of the Gaussian components.

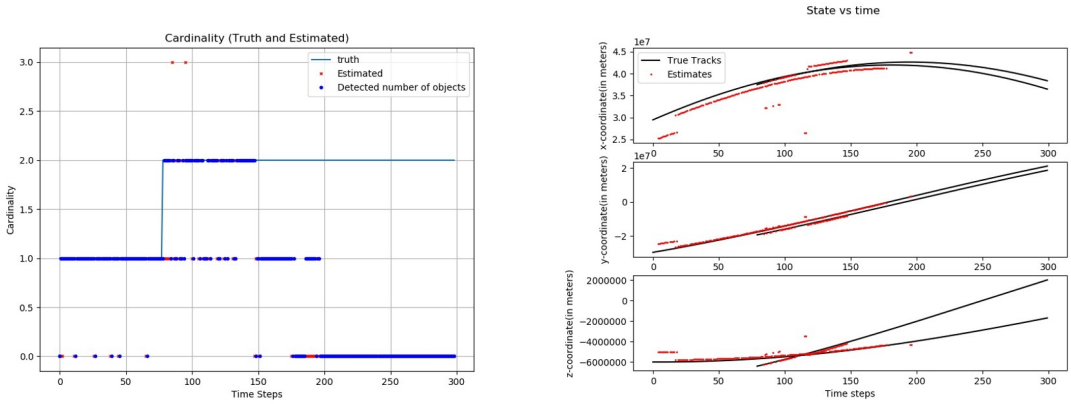
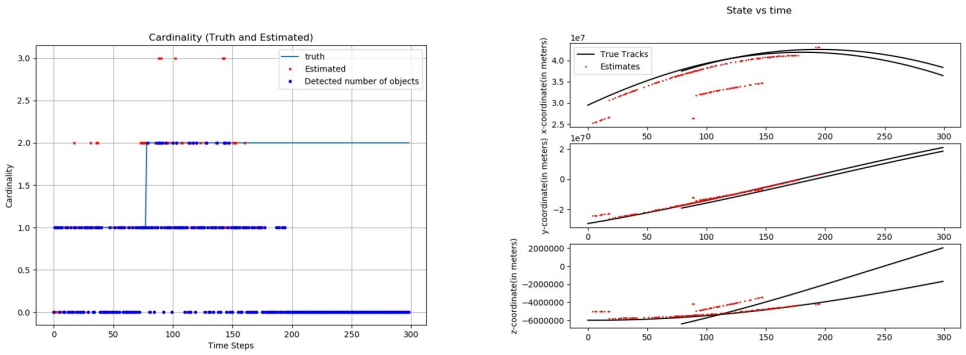


Figure 7.9. For object ground-based tracking scenario, with restricted FOV, p_D=0.9, with clutter and object birth: a) Cardinality as a function of time, blue line is the truth, in red the estimates and bright blue point is the number of target generated measurements at that time; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z

Figure 7.10 shows the results for the same case but with a lower p_D, i.e., $p_D = 0.5$. The estimates follow a similar trend as before but at a lower efficiency since there are lesser target generated measurements. The objects are regularly underestimated. This can be seen in the cardinality plot where the estimates routinely overshoot the actual number of objects. Also, the state estimates are also less accurate (have greater uncertainty) since measurements are not continuously available for the filter to update the state. This can be seen in part (b), in the x coordinate estimate of object 2 (time steps: 80-150), where the deviation from truth is comparatively higher than in part (b) of figure 7.9.



a) Cardinality Plot b) State Estimates

Figure 7.10. For object ground-based tracking scenario, with restricted FOV, p_D=0.5, with clutter and object birth: a) Cardinality as a function of time, blue line is the truth, in red the estimates and bright blue point is the number of target generated measurements at that time; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z

7.2 State Dependent p_D

As was discussed in Chapter 1, p_D should be modeled as a state-dependent quantity in the space environment. This section shows the application of the method shown in section 6.3 where p_D is modeled as a state-dependent quantity. It must be noted that there is no model mismatch considered here. This section shows the application of the p_D model in the GM-PHD filter.This simulation is run for 300 time steps and only one object is considered. The object's time in FOV and its orbital elements are given in figure 7.11.The observer is assumed to be on the surface of the Earth at a $-91.44°$ longitude and $-7.15°$ latitude and measurements are taken at every time step and consist of topocentric angles and angle-rates $(\alpha, \delta, \dot{\alpha}, \dot{\delta})$. The FOV is assumed to be $4° \times 4°$. It is assumed that the clutter poisson parameter is 4. Probability of survival is constant and is assumed to be equal to 0.99. p_D is modeled after the CCD equation (equation 6.2) and there is no model mismatch i.e., p_D for the truth and p_D used in the filter are the same. The object is assumed to be

spherical with 1 m^2 area. For the birth model, the design parameters are assumed as $\sigma_\rho = 2750, \sigma_{\dot{\rho}} = 1.2$ (see section 5.4.2).

Targets	Time steps in FOV	Semi-major axis (km)	eccentricity	Inclination (radians)	Longitude of ascending node (radians)	Argument of perigee (radians)	True Anomaly (radians)
Object 1	0-177,185-196	42164.45	0.00027	0.14278	0.80659	4.50221	0.18799

Figure 7.11. Results for state dependent p_D

In figure 7.12, part (c) shows the cardinality estimate and also the number of target generated measurements at each time, part (b) shows the state estimates. It can be seen that initially no object is tracked. This is because of the lack of target generated measurements. As the measurements are available regularly the filter responds efficiently similar to the cases in previous sections. Also, in figure 7.12, part (a) shows the estimated p_D for the object at each time step. It can be seen that estimated p_D is comparatively lower than the truth initially and hence the estimated number of objects are higher. This is because even though the target generated measurement is available, because of the lower p_D estimate, the filter assumes the measurements to mostly consist of clutter and ends up overestimating the number of objects. The estimates become more accurate as more measurements become available. And since there are no measurements available after time step 196 (since the object is out of FOV), the p_D estimate is zero and the object is tracked until the weight of it's corresponding GM-PHD component goes less than 0.5 (weight reduces continuously because $p_S = 0.99$). It is important to note that that by using this method the object is continued to be tracked even when it is not in the FOV.

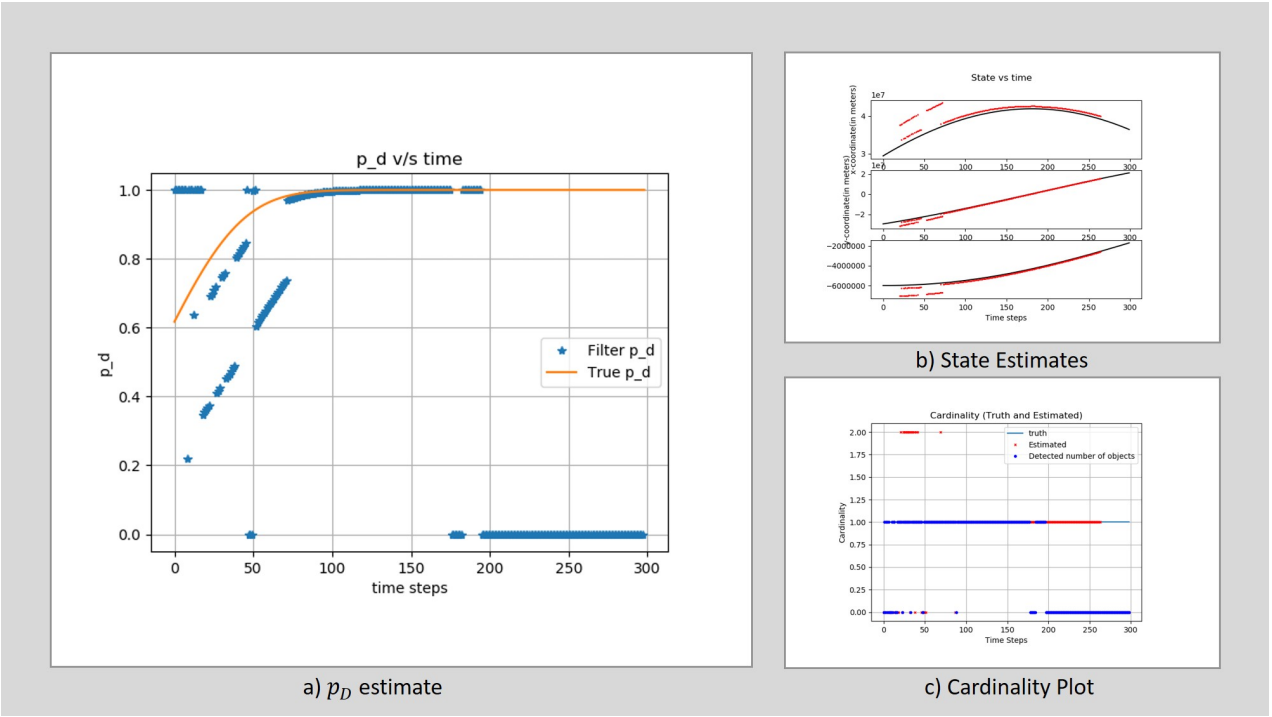


Figure 7.12. For object ground-based tracking scenario, with restricted FOV, p_D=0.5, with clutter and object birth: a) p_D as a function of time, in orange is the truth and in blue is the estimate; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z; c) Cardinality as a function of time, blue line is the truth, in red the estimates and bright blue point is the number of target generated measurements at that time

7.3 Model Mismatch

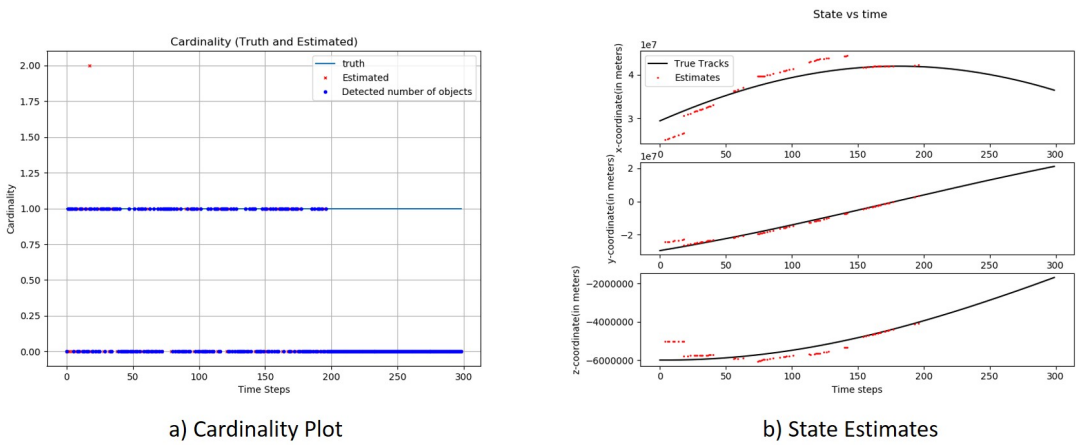
This section shows the effect of model mismatch and signifies the need to consider an uncertainty in p_D as explained in chapter 1. More precisely, a model mismatch between the true p_D and the p_D that the filter assumes. For this purpose, the GM-PHD filter is run using only a single object. This simulation is run for 300 time steps and only one object is considered. The object's time in FOV and its orbital elements are given in figure 7.11.The observer is assumed to be on the surface of the Earth at a $-91.44°$ longitude and $-7.15°$ latitude and measurements are taken at every time step

and consist of topocentric angles and angle-rates $(\alpha, \delta, \dot{\alpha}, \dot{\delta})$. The FOV is assumed to be $4^\circ \times 4^\circ$. It is assumed that the clutter poisson parameter is 4. Probability of survival is constant and is assumed to be equal to 0.99. The object is assumed to be spherical with 1 m^2 area.For the birth model, the design parameters are assumed as $\sigma_\rho = 2750, \sigma_{\dot{\rho}} = 1.2$ (see section 5.4.2).

7.3.1 Constant p_D model

The constant p_D model is assumed for the following cases for both the truth and the filter but different values are assumed for each.

The case below (Figure 7.13) assumes that the true p_D=0.5 and the filter $p_D = 0.9$. Part (a) shows the cardinality estimates and part (b) shows the state estimates. From part (a), it can be seen that the target generated measurements are not continuously available due to low p_D. Also, the number of targets are routinely underestimated. The results are as expected i.e, if the filter assumes a higher p_D than the true p_D, then filter assumes that the measurements surely contain the measurements from the targets. However, the measurements are mostly clutter and hence the filter tends to underestimate the number of targets (lack of quality measurements). It can be seen from part (b) that the state estimates, especially the x and z coordinates, have high deviation from the truth. The average error in x-coordinate is in the order of 10^3 km and in z-coordinate its in the order of 10^2 km. This is because once the track is initialized, there aren't enough measurements to reduce the uncertainty. Also, the track is initiated multiple times due to the model mismatch. The reason is that the track is cutoff because there aren't sufficient measurements (consequence of low true p_D) and also it is not continued by the $1-p_D$ term because of the high filter p_D. Also, since the object is out of FOV after the 200th time step, the estimate and the truth will be same since there are no incoming measurements. Hence, this model mismatch clearly affects the performance of the filter.



a) Cardinality Plot b) State Estimates

Figure 7.13. For object ground-based tracking scenario,constant p_D model with True p_D=0.5 and filter $p_D = 0.9$, with clutter and object birth: a) Cardinality as a function of time, blue line is the truth, in red the estimates and bright blue point is the number of target generated measurements at that time; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z

In the following case (figure 7.14), the p_D values are reversed. Since the filter assumes a lower p_D value than the true p_D value, it assumes that the measurements contain less number of target measurements. However, since true p_D is high (0.9), target measurements are surely there and because of this the filter overestimates the number of targets. This can be seen in the cardinality plot (Figure 7.14, (a)) where the estimates show 2 objects at the same state instead of one. It can also be seen from the state estimate plots (part (b)) that, there is high deviation in the x, z coordinate estimates. The max deviation in x is in the order of 10^4 km and in z its in the order of 10^3 km. Hence, the filter performance is affected in this case too.

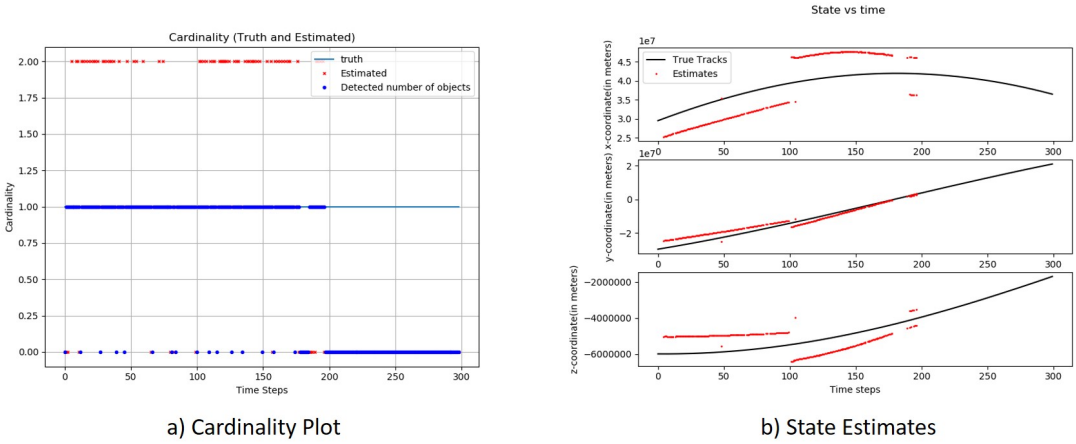


Figure 7.14. For object ground-based tracking scenario,constant p_D model with True p_D=0.9 and filter $p_D = 0.5$, with clutter and object birth: a) Cardinality as a function of time, blue line is the truth, in red the estimates and bright blue point is the number of target generated measurements at that time; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z

7.3.2 State dependent p_D

In this section, the state dependent p_D model is assumed for the truth whereas a constant p_D model is assumed for the filter and the results are analysed.

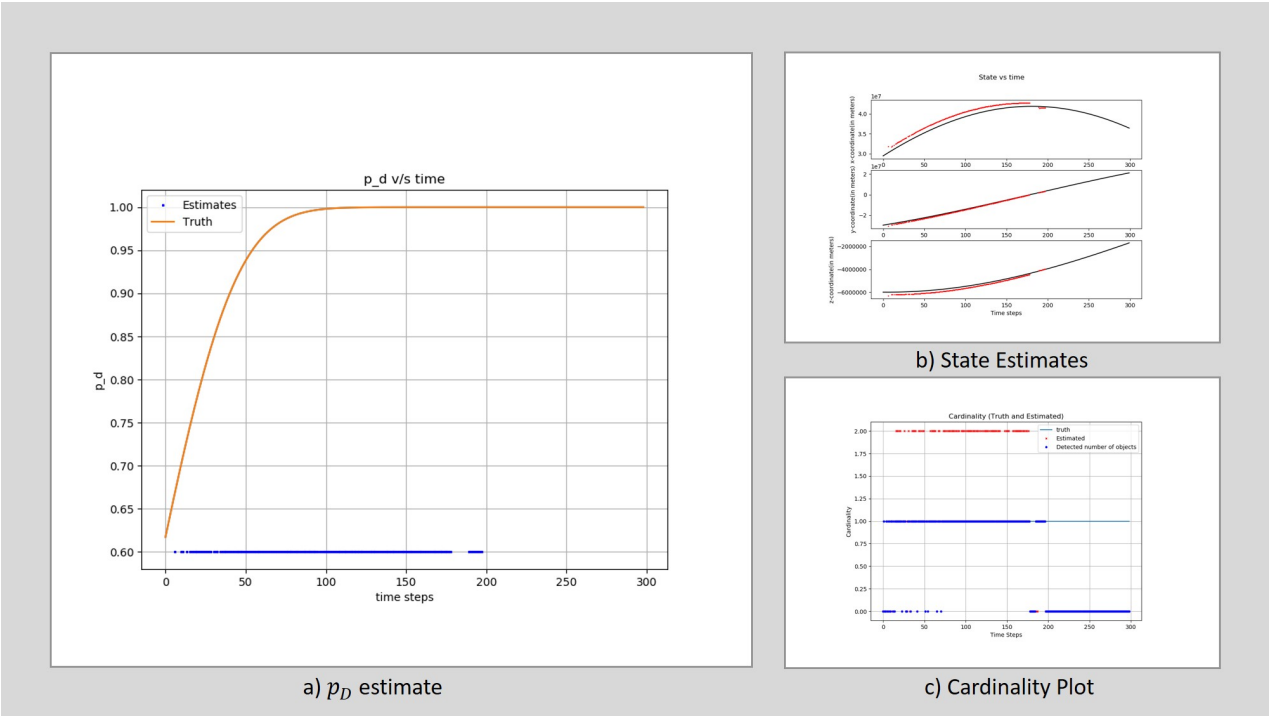


Figure 7.15. For object ground-based tracking scenario, state dependent p_D model for the truth and constant p_D model for the filter with $p_D = 0.6$, with clutter and object birth: a) p_D as a function of time, in orange is the truth and in blue is the estimate; b) the state estimates of the objects, in black the true tracks, in red the estimates displayed in position x,y,z; c) Cardinality as a function of time, blue line is the truth, in red the estimates and bright blue point is the number of target generated measurements at that time

Initially the filter p_D is closer to the true p_D and hence the estimates are closer to the truth. But as time progresses, since the true p_D is state dependent it increases whereas the filter p_D is constant at 0.6. Therefore, the filter overestimates the number of targets since the filter p_D is lower compared to the true p_D. This shows that the constant p_D assumption in the GM-PHD filter is not a good assumption in the space environment.

7.4 Uncertainty in p_D

As can be seen from the previous section a model mismatch in the probability of detection affects the performance of the GM-PHD filter. Especially, over-predicting the p_D poses a greater concern since this leads to under-predicting the number of targets and this leads to greater problems. For example, if the tracking is done to avoid the collisions, under-predicting the targets is not desired. Hence, it is necessary to incorporate an uncertainty in the probability of detection if there is one. The model described in section 6.4 is used to for this purpose. As mentioned previously, an uncertainty in Area results in the uncertainty in p_D. The efficiency of this method is established by comparing the performance of the filter which takes the uncertainty into account with the filter which doesn't take the uncertainty into account. The following paragraph describe this process.

The area (A) of the targets is assumed to be a random variable with a Gaussian Distribution. Then, a hundred samples of Area are taken and for each sample a truth (measurements) is generated, i.e., the measurements differ for each sample since a change in area results in a change in p_D which affects the observations. Clutter measurements are also included. The filter is run using three models of p_D for 100 samples (100 sets of measurements) and the results are stored. For the first model, namely "M1", the state dependent p_D model (see section 6.3) is used where the Area is assumed to be constant and equal to the mean of the Gaussian Distribution. The model M1 doesn't consider the uncertainty in the probability of detection. For the second model "M2", the state dependent p_D model is used but this time the value of p_D is taken as a weighted average i.e., 100 samples of areas are drawn (from the Gaussian distribution) and p_D is calculated for each sample and a histogram with 10 bins (0-0.1,0.1-0.2,...,0.9-1) is created, and, a weighted average using the centers of the bins is calculated. The third model "M3" is described in section 6.4. As mentioned above, the filter is applied on the 100 sets of measurements and the results for the state estimates and the cardinality (number of objects) estimates at each time step

are stored. At each time step, an error is calculated using the Optimal Subpattern Assignment metric, also known as OSPA (refer [51]). The error is then averaged at each time step for the 100 samples for all the models and the error is plotted. The model "M3" can be validated if the error for "M3" is less than the error for "M1". Note that the model "M2" uses a Monte-Carlo methodology which is a valid model, but it is computationally expensive. This thesis provides an alternative method (M3) which is closer in efficiency to "M2" but is computationally less expensive.

7.4.1 Results for GEO objects

This section shows the error comparison for the models (M1,M2,M3) for a test case where two GEO objects are tracked with a restricted FOV. For all of the following results the clutter and the observer properties are same as the ones described in section 7.1.2. It is assumed that the target's area has a Gaussian Distribution with mean 1 and a standard deviation of 0.3.

The following figure show the results when only object is considered (object 1 in section 7.1.2).

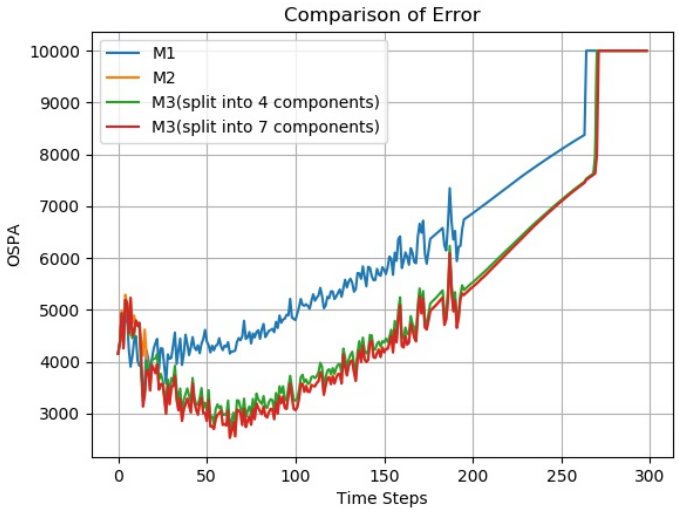


Figure 7.16. Comparison of OSPA metric errors (Clutter included)

As can be seen from figure 7.16, the error in M1 is the highest since it doesn't take the uncertainty into account and ends up either overestimating or underestimating the number of targets as described in section 7.3. The model M3 performs efficiently as compared to the model M1 and is as good as the model M2 which is a Monte Carlo analysis. The errors are identical at the end since the object is out of the FOV after the 200th time step and hence the three filter models for estimating p_D work the same.

The next figure shows the result when two objects are considered (object 1, object 2 in section 7.1.2). Clutter was not considered in this case. However, the trend will be similar even if clutter is considered (figure 7.18).

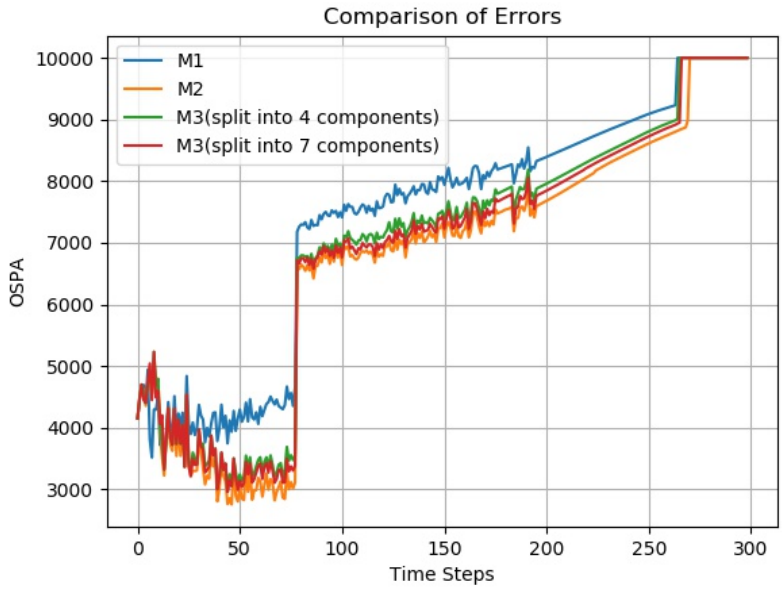


Figure 7.17. Comparison of OSPA metric errors (Clutter not included)

The same conclusions can be drawn again. There is a jump around 80th time step because object 2 enters the FOV at this time step and, this leads to higher error.

From figures 7.16 and 7.18 it can be seen that the model M3 performance is similar to the performance of model M2. Also, from section 6.4.1, it can be seen that the p_D estimates obtained from model M2 and M3 are pretty close. Hence, for the rest of

this section, the results for model M2 are not shown. The following figure shows the result for two objects when clutter is considered.

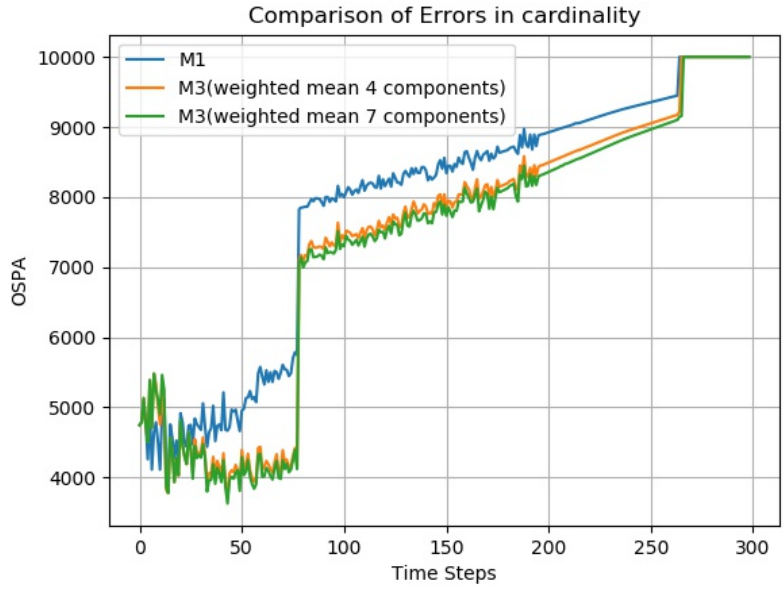
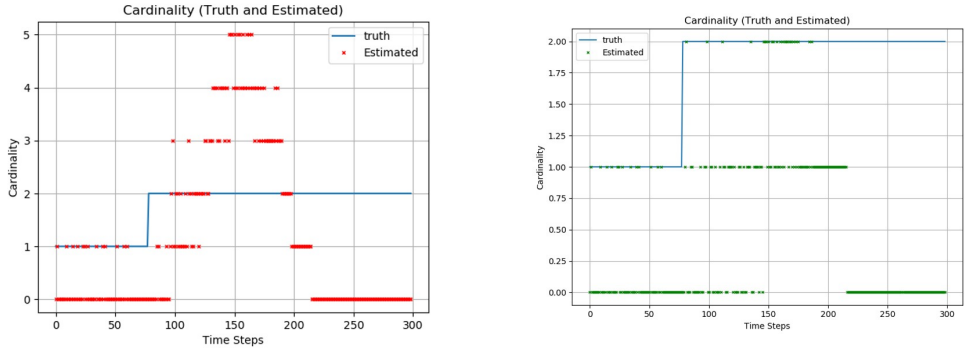


Figure 7.18. Comparison of OSPA metric errors (with Clutter)

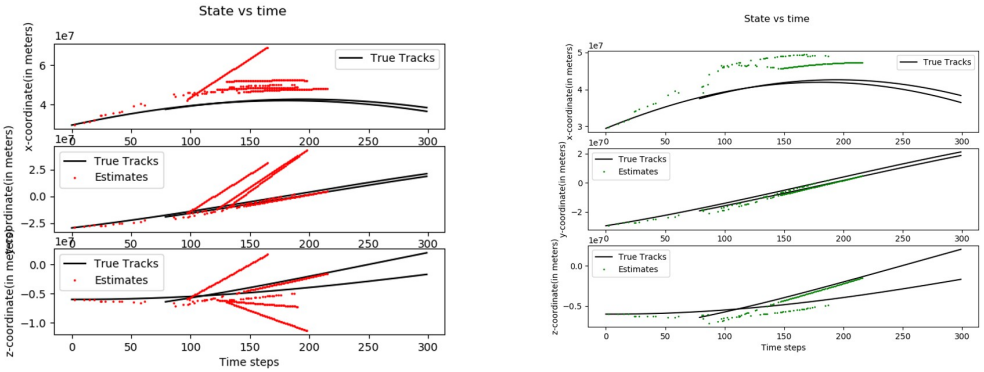
The results follow a similar trend as before. The following figures (figure 7.19, figure 7.20, figure 7.21 and figure 7.22) show the cardinality and state plots for two cases of area samples (area of objects = 0.6745, 0.9672). Part(a) shows the plots when model M1 is applied and part (b) shows the results when model M3 (7 components) is applied.

Figures 7.19 and 7.20 are obtained by assuming the area of the objects to be 0.6745. Since model M1 assumes that the area is 1, there is a model mismatch here. This can be reflected in the cardinality plots where the number of targets are overestimated because the estimated p_D is lower than the true p_D. But the model M3 models p_D as a uncertain quantity and hence the results are more accurate.



a) Model M1 b) Model M3, 7 components

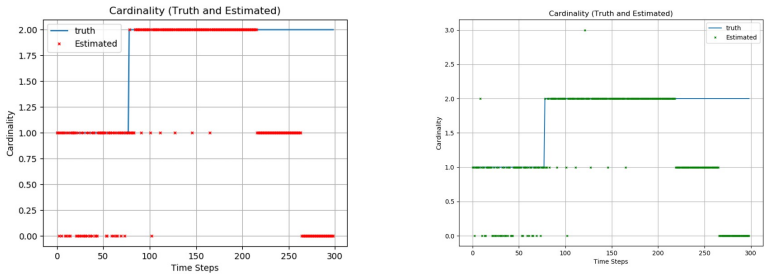
Figure 7.19. For object ground-based tracking scenario, with restricted FOV, without clutter and with object birth, area of objects = 0.6745, plots show cardinality as a function of time obtained by application of models M1 and M3, blue line is the truth, in red (Model M1) and green (Model M3) the estimates: a) p_D estimate obtained by model M1; b) p_D estimate obtained by model M3



a) Model M1 b) Model M3, 7 components

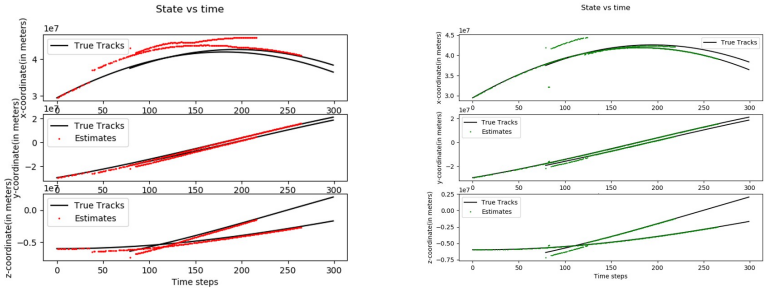
Figure 7.20. For object ground-based tracking scenario, with restricted FOV, without clutter and with object birth, area of objects = 0.6745, plots show state estimates (position x,y,z) as a function of time obtained by application of p_D models M1 and M3, black line is the truth, in red (Model M1) and green (Model M3) the estimates: a) p_D estimate obtained by model M1; b) p_D estimate obtained by model M3

Figures 7.21 and 7.22 are obtained by assuming the area of the objects to be 0.9672. In this case the performance of model M1 and model are similar because the area assumed by the model M1 is close to the true value. Hence, the p_D estimates are similar in both the models and hence the performance is similar.



a) Model M1 b) Model M3, 7 components

Figure 7.21. For object ground-based tracking scenario, with restricted FOV, without clutter and with object birth, area of objects = 0.9672, plots show cardinality as a function of time obtained by application of models M1 and M3, blue line is the truth, in red (Model M1) and green (Model M3) the estimates: a) p_D estimate obtained by model M1; b) p_D estimate obtained by model M3



a) Model M1 b) Model M3, 7 components

Figure 7.22. For object ground-based tracking scenario, with restricted FOV, without clutter and with object birth, area of objects = 0.9672, plots show state estimates (position x,y,z) as a function of time obtained by application of p_D models M1 and M3, black line is the truth, in red (Model M1) and green (Model M3) the estimates: a) p_D estimate obtained by model M1; b) p_D estimate obtained by model M3

The next section shows the results when LEO objects are considered.

7.4.2 Results for LEO objects

In this section, the targets are taken to be in LEO. This simulation is run for 135 time steps with a step size of 1 minute. The objects orbital elements are shown in 7.1.2. The orbit elements are shown in figure 7.23. The observer is assumed to be on the surface of the Earth at a $-91.44°$ longitude and $-7.15°$ latitude and measurements are taken at every time step and consist of topocentric angles and angle-rates $(\alpha, \delta, \dot{\alpha}, \dot{\delta})$. The FOV model used here is described in 5.3.1 and is assumed to be $16° \times 16°$. It is assumed that the clutter Poisson parameter is 4. Probability of survival is constant and is assumed to be equal to 0.99. For the birth model, the design parameters are changed to $\sigma_\rho = 200, \sigma_{\dot{\rho}} = 0.5$ (see section 5.4.2) since the Admissible region shrinks as compared to GEO objects and hence smaller design components have to be taken to have necessary number of GM components to fill out the Admissible Region.

Targets	Semi-major axis (km)	eccentricity	Inclination (radians)	Longitude of ascending node (radians)	Argument of perigee (radians)	True Anomaly (radians)
Object 1	7378.323	0.01034	1.44877	4.10412	4.01713	1.39263
Object 2	7849.591	0.05620	1.74179	3.53892	1.92272	1.97052
Object 3	7184.974	0.00870	1.45063	4.25730	2.65441	2.60532

Figure 7.23. Orbital elements of the targets

The simulation is run for The FOV restriction for LEO objects was described in section 5.3.1. The clutter is modeled similarly as described in the previous section. The simulation is run for 135 time steps with each time step equal to 60 seconds. The design parameters for the birth model are changed to $\sigma_\rho = 200, \sigma_{\dot{\rho}} = 0.5$ since the Admissible region shrinks as compared to GEO objects and hence smaller design components have to be taken to have necessary number of GM components to fill

out the Admissible Region. The following figure shows the FOV movement for this scenario. Figure 7.25 displays the results for this case.

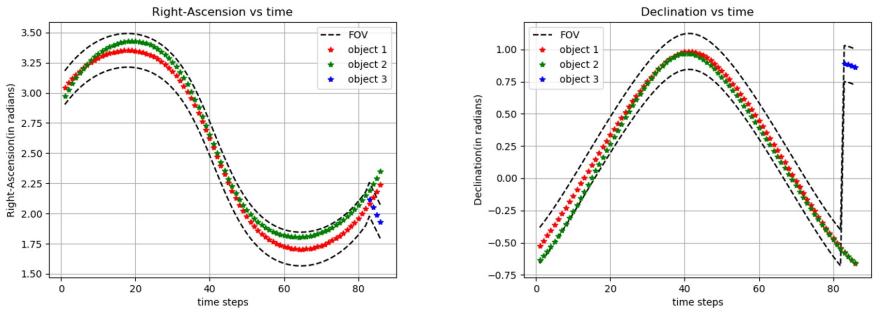


Figure 7.24. Field of View movement

Figure 7.25. Comparison of Error

It can be seen from figure 7.25 that the model M3 works better compared to the model M1. However, the error contrast at the end of the time interval (after the 40th time step) is not very high because the true p_D of the LEO objects used in the simulation is closer to one in the time interval considered (see figure 7.26) i.e., if the objects' area is equal to the mean of the Gaussian distribution (it is 1 here), their p_D

is mostly closer to one. Hence, a variation in area won't affect the p_D drastically and both the models will give similar p_D estimates. The following figure shows the true p_D for the objects at the mean area.

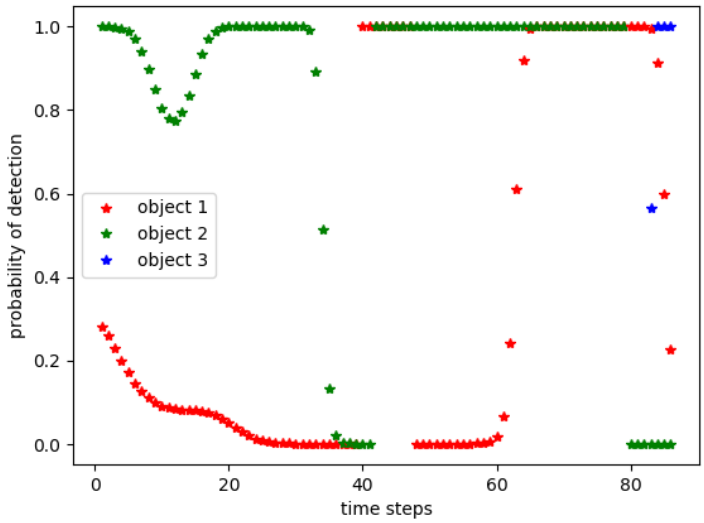
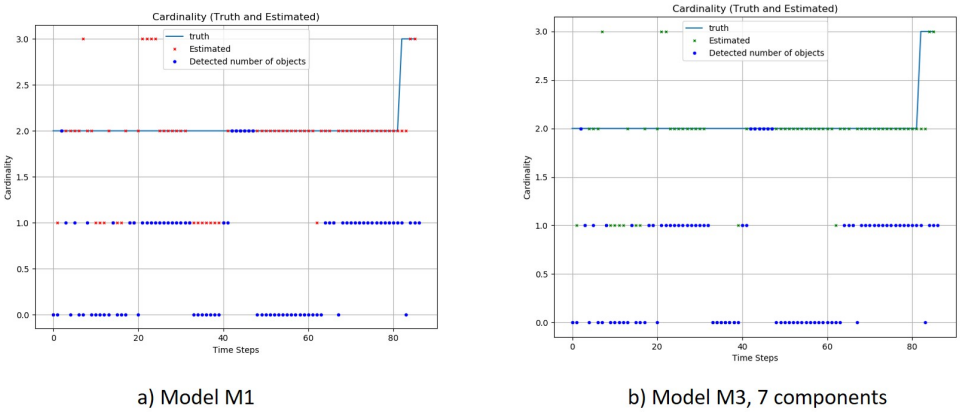


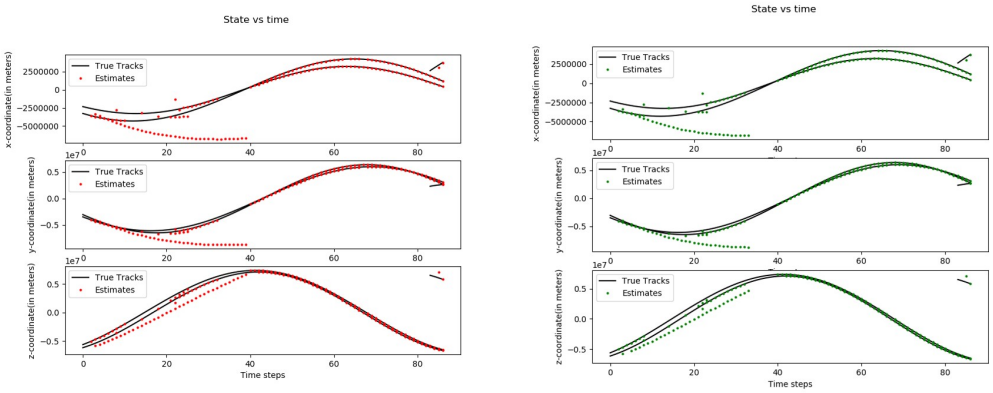
Figure 7.26. True p_D of the objects at the mean area

The following figures (figure 7.27, figure 7.28, figure 7.29 and figure 7.30) show the cardinality and state plots for two cases of area samples (area of objects = 0.6745, 0.9672). Part(a) shows the plots when model M1 is applied and part (b) shows the results when model M3 (7 components) is applied. Figures 7.27, 7.28 show the results when the area of the objects is 0.6745 whereas figures 7.29, 7.30 show results when the area is 0.9762. It can be seen from the figures that the performance of model M1 and model M3 is similar. This is because the true p_D (see figure 7.26) of the objects at the mean area is closer to one. Hence, unless the area of the objects is much lesser than 1, both the models will give the same p_D estimates. This is one of the reason why 100 cases of area are considered since less number of area samples don't depict the complete picture.



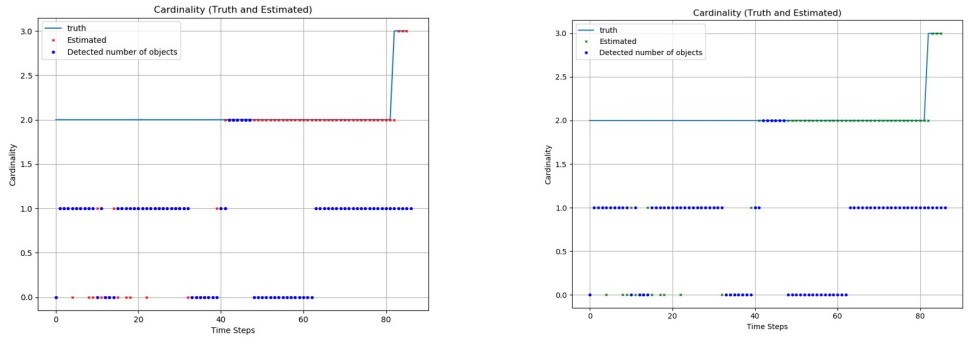
a) Model M1 b) Model M3, 7 components

Figure 7.27. For object ground-based tracking scenario, with restricted FOV, without clutter and with object birth, area of objects = 0.6745, plots show cardinality as a function of time obtained by application of models M1 and M3, blue line is the truth, in red (Model M1) and green (Model M3) the estimates: a) p_D estimate obtained by model M1; b) p_D estimate obtained by model M3



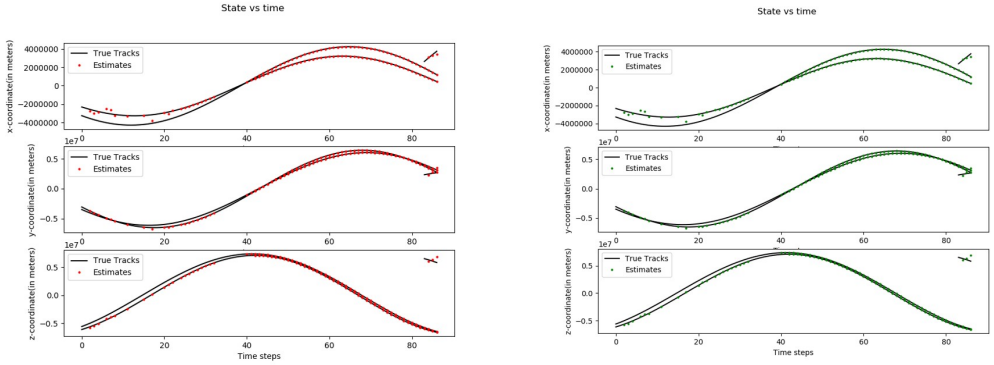
a) Model M1 b) Model M3, 7 components

Figure 7.28. For object ground-based tracking scenario, with restricted FOV, without clutter and with object birth, area of objects = 0.6745, plots show state estimates (position x,y,z) as a function of time obtained by application of p_D models M1 and M3, black line is the truth, in red (Model M1) and green (Model M3) the estimates: a) p_D estimate obtained by model M1; b) p_D estimate obtained by model M3



a) Model M1 b) Model M3, 7 components

Figure 7.29. For object ground-based tracking scenario, with restricted FOV, without clutter and with object birth, area of objects = 0.9672, plots show cardinality as a function of time obtained by application of models M1 and M3, blue line is the truth, in red (Model M1) and green (Model M3) the estimates: a) p_D estimate obtained by model M1; b) p_D estimate obtained by model M3



a) Model M1 b) Model M3, 7 components

Figure 7.30. For object ground-based tracking scenario, with restricted FOV, without clutter and with object birth, area of objects = 0.9672, plots show state estimates (position x,y,z) as a function of time obtained by application of p_D models M1 and M3, black line is the truth, in red (Model M1) and green (Model M3) the estimates: a) p_D estimate obtained by model M1; b) p_D estimate obtained by model M3

Therefore, it is seen that when there is an uncertainty in estimating the area of the targets, the filter which doesn't include the p_D uncertainty (model M1) has a lesser efficiency than the filters which take p_D uncertainty into account (models M2,M3). Also, model M3, i.e., the proposed method works as efficiently as the model M2, which implements a Monte-Carlo methodology.

8. SUMMARY AND RECOMMENDATIONS

8.1 Conclusions

Multi-target tracking is an integral part of Space Situational Awareness (SSA). A method established in other tracking contexts, and also applied to SSA, is the Gaussian Mixture Probability Hypothesis Density (GM-PHD) filter. An important component of this algorithm is the probability of detection (p_D) which determines how likely the target generates a measurement. GM-PHD filter, in its original form, models p_D as a state-independent quantity i.e., a constant. It is seen that in SSA, p_D is a state-dependent quantity and also has some uncertainty associated with it. This research successfully incorporated uncertainty in the p_D and also modeled p_D as a state (and therefore) time dependent within the existing GM-PHD framework.
For the application in a SSA system, birth was modeled via admissible regions, and an extended Kalman filter was used for orbit propagation and update. The objects were assumed to follow a two-body motion model and for all the test cases, an optical ground-based single sensor scenario has been used. A simulation scenario in SSA was designed and the effects of model mismatch in p_D in the GM-PHD filter was investigated. It is seen that the filter's efficiency is reduced if there is a model mismatch in p_D i.e., if the true p_D is state-dependent (as is the case in SSA) and if the filter assumes p_D as a constant, the error in the estimates are not negligible. The filter either overestimates or underestimates the number of targets in the scene depending on the mismatch. Also, p_D was not modeled as a state-dependent quantity in any of the previous research work done in SSA. Hence, a variation of the GM-PHD filter which models p_D as state-dependent and also which incorporates p_D uncertainty was developed. The uncertainty in p_D was created from an uncertainty in the reflecting area of objects. The area of objects was assumed to be a Gaussian random variable.

Since the relationship between p_D and area is non-linear, p_D is linearly approximated as a Gaussian Mixture.

The proposed method to incorporate uncertainty was validated with simulations and it was seen that it performs better than the filter which doesn't accommodate uncertainty in p_D.

8.2 Recommendations

There are still many variations that can be made to the GM-PHD filter to use it in a typical SSA scenario. The proposed method can also be extended to the CPHD filter which propagates the PHD as well as the cardinality density. As mentioned before, the p_D model used in this work is modeled as a function of the sensor properties and the amount of light reflected by the object. For the Earth-Orbiting objects, especially the objects in LEO, the shadow of the Earth also has to be considered in modeling p_D since without the irradiation from the Sun there would be no reflection. Also, the amount of reflected light from the objects also depends on the material of the reflecting surface which is also very difficult to estimate without a complete knowledge of the object. Hence, there is an uncertainty in the estimating the reflectivity of the material too. A future work could look into exploring a new model for p_D which considers these factors.

All the objects considered in the simulations are assumed to be under a Two-Body motion. However, in reality, space objects are influenced not only by Earth, but also experience perturbations from gravitational pull of the Moon and Sun and solar radiation pressure. These perturbations would influence the measurements obtained and hence, to account for this, the motion model has to be changed.

As was discussed above, the Admissible Region (AR) was used to design the Birth model for the GM-PHD filter. The AR considered here is an unconstrained one, i.e., the region lists all the possible Earth Captured orbits for a particular measurement. However, in some cases, there is partial knowledge available on the objects to be

tracked. For example, consider the case where some specific GEO objects are being tracked. Then, further constraints can be applied to the AR which reduces the computational burden of the filter. In the case of GEO objects, a semi-major axis constraint or a eccentricity constraint can be applied. Constrained Admissible region methods are detailed in [43] and [44]. Future work can use the constrained admissible regions for designing the birth model.

REFERENCES

REFERENCES

[1] B-N Vo and W-K Ma. The gaussian mixture probability hypothesis density filter. *IEEE Transactions on signal processing*, 54(11):4091–4104, 2006.

[2] T.S.Kelso. *SATCAT Boxscore*, July 2019.

[3] Loren Grush. *SpaceX just launched two of its space internet satellites the first of nearly 12,000*, February 2018.

[4] GW Pulford. Taxonomy of multiple target tracking methods. *IEE Proceedings-Radar, Sonar and Navigation*, 152(5):291–304, 2005.

[5] Samuel S Blackman. Multiple hypothesis tracking for multiple target tracking. *IEEE Aerospace and Electronic Systems Magazine*, 19(1):5–18, 2004.

[6] Jeffrey K Uhlmann. Algorithms for multiple-target tracking. *American Scientist*, 80(2):128–141, 1992.

[7] Jeffrey K Uhlmann, M Liggins, D Hall, and J Llinas. An introduction to the combinatorics of optimal and approximate data association. In *Handbook of Multisensor Data Fusion: Theory and Practice*. CRC Press, 2008.

[8] Aubrey B Poore, Suihua Lu, and Brian J Suchomel. Data association using multiple frame assignments. *Handbook of Multisensor Data Fusion, CRC Press LLC*, 2009.

[9] Yaakov Bar-Shalom and Xiao-Rong Li. Estimation and tracking- principles, techniques, and software. *Norwood, MA: Artech House, Inc, 1993.*, 1993.

[10] Samuel S Blackman. Multiple-target tracking with radar applications. *Dedham, MA, Artech House, Inc., 1986, 463 p.*, 1986.

[11] S Blackman and R Popoli. Design and analysis of modern tracking systems, ser. *Artech House radar library. Artech House*, 1999.

[12] Ronald PS Mahler. *Statistical multisource-multitarget information fusion*. Artech House, Inc., 2007.

[13] Thomas Fortmann, Yaakov Bar-Shalom, and Molly Scheffe. Sonar tracking of multiple targets using joint probabilistic data association. *IEEE journal of Oceanic Engineering*, 8(3):173–184, 1983.

[14] C Frueh and T Schildknecht. Object image linking of objects in near earth orbits in the presence of cosmics. *Advances in Space Research*, 49(3):594–602, 2012.

[15] Carolin Früh. *Identification of Space Debris*. Shaker, 2011.

[16] Thomas Kelecy, Moriba Jah, and Kyle DeMars. Application of a multiple hypothesis filter to near geo high area-to-mass ratio space objects state estimation. *Acta Astronautica*, 81(2):435–444, 2012.

[17] Navraj Singh, Joshua T Horwood, Jeffrey M Aristoff, Aubrey B Poore, Carolyn Sheaff, and Moriba K Jah. Multiple hypothesis tracking (mht) for space surveillance: Results and simulation studies. Technical report, AIR FORCE RESEARCH LAB KIRTLAND AFB NM SPACE VEHICLES DIRECTORATE, 2013.

[18] Kyle J DeMars, Islam I Hussein, Carolin Frueh, Moriba K Jah, and R Scott Erwin. Multiple-object space surveillance tracking using finite-set statistics. *Journal of Guidance, Control, and Dynamics*, 38(9):1741–1756, 2015.

[19] Steven Gehly, Brandon Jones, and Penina Axelrad. An aegis-cphd filter to maintain custody of geo space objects with limited tracking data. In *Proceedings of the 2014 Advanced Maui Optical and Space Surveillance Technologies Conference*, pages 25–34, 2014.

[20] Brandon A Jones, Steve Gehly, and Penina Axelrad. Measurement-based birth model for a space object cardinalized probability hypothesis density filter. In *AIAA/AAS Astrodynamics Specialist Conference*, page 4311, 2014.

[21] Brandon A Jones and Ba-Ngu Vo. A labeled multi-bernoulli filter for space object tracking. In *Proceedings of the 2014 AAS/AIAA Spaceflight Mechanics Meeting, Santa Fe, NM, USA*, pages 26–30, 2014.

[22] Yang Cheng, Carolin Früh, and Kyle J DeMars. Comparisons of phd filter and cphd filter for space object tracking. 2014.

[23] Y Cheng, KJ DeMars, C Früh, and MK Jah. Gaussian mixture phd filter for space object tracking. In *Proceedings of the AAS/AIAA Space Flight Mechanics Meeting, ser. Advances in the Astronautical Sciences*, 2013.

[24] C Frueh. Modeling impacts on space situational awareness phd filter tracking. *CMES-COMPUTER MODELING IN ENGINEERING & SCIENCES*, 111(2):171–201, 2016.

[25] Yaakov Bar-Shalom, Thomas E Fortmann, and Peter G Cable. Tracking and data association, 1990.

[26] Daniel Clark, Ba-Ngu Vo, and Judith Bell. Gm-phd filter multitarget tracking in sonar images. In *Signal Processing, Sensor Fusion, and Target Recognition XV*, volume 6235, page 62350R. International Society for Optics and Photonics, 2006.

[27] Ya-Dong Wang, Jian-Kang Wu, Weimin Huang, and Ashraf A Kassim. Gaussian mixture probability hypothesis density for visual people racking. In *2007 10th International Conference on Information Fusion*, pages 1–6. IEEE, 2007.

[28] Yun Zhang, Huilin Mu, Yicheng Jiang, Chang Ding, and Yong Wang. Moving target tracking based on improved gmphd filter in circular sar system. *IEEE Geoscience and Remote Sensing Letters*, 2018.

[29] Carolin Früh and Moriba K Jah. Detection probability of earth orbiting objects using optical sensors. *Advances in the Astronautical Sciences, Hilton Head*, 2014.

[30] HW Sorenson. Recursive estimation for nonlinear dynamic systems. *Bayesian analysis of time series and dynamic models*, 94:127–165, 1988.

[31] Fred Daum. Nonlinear filters: beyond the kalman filter. *IEEE Aerospace and Electronic Systems Magazine*, 20(8):57–69, 2005.

[32] Andrew H Jazwinski. *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[33] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.

[34] Ronald Mahler. Multitarget moments and their application to multitarget tracking. Technical report, LOCKHEED MARTIN TACTICAL DEFENSE SYSTEMS-EAGAN ST PAUL MN, 2001.

[35] Irwin R Goodman, Ronald P Mahler, and Hung T Nguyen. *Mathematics of data fusion*, volume 37. Springer Science & Business Media, 2013.

[36] Ronald Mahler. Multitarget sensor management of dispersed mobile sensors. In *Theory and algorithms for cooperative systems*, pages 239–310. World Scientific, 2004.

[37] Wing-Kin Ma, Ba-Ngu Vo, Sumeetpal S Singh, and Adrian Baddeley. Tracking an unknown time-varying number of speakers using tdoa measurements: A random finite set approach. *IEEE Transactions on Signal Processing*, 54(9):3291–3304, 2006.

[38] Ronald PS Mahler. Multitarget bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic systems*, 39(4):1152–1178, 2003.

[39] Daniel E Clark, Kusha Panta, and Ba-Ngu Vo. The gm-phd filter multiple target tracker. In *2006 9th International Conference on Information Fusion*, pages 1–8. IEEE, 2006.

[40] Daniel Clark and Ba-Ngu Vo. Convergence analysis of the gaussian mixture phd filter. *IEEE Transactions on Signal Processing*, 55(4):1204–1212, 2007.

[41] Prasanta Chandra Mahalanobis. On the generalized distance in statistics. National Institute of Science of India, 1936.

[42] Carolin Frueh. Lecture notes of the course : Space traffic management,purdue university, Fall 2018.

[43] Andrea Milani, Giovanni F Gronchi, Mattia deMichieli Vitturi, and Zoran Knežević. Orbit determination with very short arcs. i admissible regions. *Celestial Mechanics and Dynamical Astronomy*, 90(1-2):57–85, 2004.

[44] Giacomo Tommei, Andrea Milani, and Alessandro Rossi. Orbit determination of space debris: admissible regions. *Celestial Mechanics and Dynamical Astronomy*, 97(4):289–304, 2007.

[45] Kyle J DeMars and Moriba K Jah. Probabilistic initial orbit determination using gaussian mixture models. *Journal of Guidance, Control, and Dynamics*, 36(5):1324–1335, 2013.

[46] Simon Julier and Jeffrey K. Uhlmann. A general method for approximating nonlinear transformations of probability distributions. Technical report, 1996.

[47] Francois Sanson and Carolin Frueh. Noise quantification in optical observations of resident space objects for probability of detection and likelihood. In *AAS/AIAA Astrodynamic Specialist Conference, Vail, CO*, pages 15–634, 2015.

[48] Gustaf Hendeby and Rickard Karlsson. Gaussian mixture phd filtering with variable probability of detection. In *17th International Conference on Information Fusion (FUSION)*, pages 1–7. IEEE, 2014.

[49] Kyle J DeMars, Robert H Bishop, and Moriba K Jah. Entropy-based approach for uncertainty propagation of nonlinear dynamical systems. *Journal of Guidance, Control, and Dynamics*, 36(4):1047–1057, 2013.

[50] Marco F Huber, Tim Bailey, Hugh Durrant-Whyte, and Uwe D Hanebeck. On entropy approximation for gaussian mixture random vectors. In *Multisensor Fusion and Integration for Intelligent Systems, 2008. MFI 2008. IEEE International Conference on*, pages 181–188. IEEE, 2008.

[51] Dominic Schuhmacher, Ba-Tuong Vo, and Ba-Ngu Vo. A consistent metric for performance evaluation of multi-object filters. *IEEE Transactions on Signal Processing*, 56(8):3447–3457, 2008.