

**Estimation of Geosynchronous Space Objects Using Finite
Set Statistics Filtering Methods**

by

Steven Gehly

B.S., Aerospace Engineering, University of Southern California, 2006

M.S., Astronautics Engineering, University of Southern California, 2009

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Aerospace Engineering Sciences

2016

This thesis entitled:
Estimation of Geosynchronous Space Objects Using Finite Set Statistics Filtering Methods
written by Steven Gehly
has been approved for the Department of Aerospace Engineering Sciences

Dr. Penina Axelrad

Dr. Brandon Jones

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Gehly, Steven (Ph.D., Aerospace Engineering Sciences)

Estimation of Geosynchronous Space Objects Using Finite Set Statistics Filtering Methods

Thesis directed by Dr. Penina Axelrad

The use of near Earth space has increased dramatically in the past few decades, and operational satellites are an integral part of modern society. The increased presence in space has led to an increase in the amount of orbital debris, which poses a growing threat to current and future space missions. Characterization of the debris environment is crucial to our continued use of high value orbit regimes such as the geosynchronous (GEO) belt. Objects in GEO pose unique challenges, by virtue of being densely spaced and tracked by a limited number of sensors in short observation windows. This research examines the use of a new class of multitarget filters to approach the problem of orbit determination for the large number of objects present. The filters make use of a recently developed mathematical toolbox derived from point process theory known as Finite Set Statistics (FISST). Details of implementing FISST-derived filters are discussed, and a qualitative and quantitative comparison between FISST and traditional multitarget estimators demonstrates the suitability of the new methods for space object estimation. Specific challenges in the areas of sensor allocation and initial orbit determination are addressed in the framework. The sensor allocation scheme makes use of information gain functionals as formulated for FISST to efficiently collect measurements on the full multitarget system. Results from a simulated network of three ground stations tracking a large catalog of geosynchronous objects demonstrate improved performance as compared to simpler, non-information theoretic tasking schemes. Further studies incorporate an initial orbit determination technique to initiate new tracks in the multitarget filter. Together with a sensor allocation scheme designed to search for new targets and maintain knowledge of the existing catalog, the method comprises a solution to the search-detect-track problem. Simulation results for a single sensor case show that the problem can be solved for multiple objects with no *a priori* information, even in the presence of missed detections and false measurements.

Collectively, this research seeks to advance the capabilities of FISST-derived filters for use in the estimation of geosynchronous space objects; additional directions for future research are presented in the conclusion.

“Art is partly communication, but only partly. The rest is discovery.”

-William Golding

“Personally, I liked the university.”

-Dr. Raymond Stantz

Acknowledgements

I would like to thank the members of my committee, Dr. Jay McMahon, Dr. Nisar Ahmed, Dr. Gregory Beylkin, and especially Dr. Brandon Jones and Dr. Penina Axelrad for their patience and guidance with this project. I would also like to thank Dr. Daniel Clark for preparing a class on the mathematical background of Finite Set Statistics which was instrumental to my understanding of the subject. Additionally, I would like to acknowledge my fellow classmates, Daniel Bryant, Nicola Baresi, In-Kwan Park, Illán García Amor, and Nikhil Shetty for their work in deriving the CPHD spawning model included in the appendix.

Partial funding for this project was provided by Air Force SBIR Contract Number FA9451-14-M-0182 awarded to Orbit Logic Incorporated. I wish to thank Alex Herz, Ella Herz, Doug George, and Ian Zeigler of Orbit Logic for their conversations and contributions to this work. I additionally wish to acknowledge Dr. Ben Bradley for his contribution of software to generate the catalog of objects from TLE data and to perform frame rotations per the latest IERS standard, and Dr. Brandon Jones and Dr. Keric Hill for use of the TurboProp software for orbit prediction. This document has been reviewed and approved for public release by the Air Force Research Laboratory Space Vehicles Directorate.

Contents

Chapter	
1	Introduction 1
1.1	Motivation 1
1.2	Challenges in Space Situational Awareness 3
1.3	Overview of Research 4
1.3.1	Comparison of Multitarget Filtering Methods 6
1.3.2	Information Theoretic Sensor Allocation 7
1.3.3	Initial Orbit Determination and Follow-On Tracking 8
1.3.4	Contributions of Research 8
1.4	Thesis Overview 9
2	Comparison of Multitarget Filtering Methods 11
2.1	Review of Single Target Filtering 12
2.1.1	AEGIS UKF 13
2.2	Conventional Multitarget Filtering 18
2.2.1	Multiple Hypothesis Tracking 19
2.2.2	Joint Probabilistic Data Association 23
2.3	Finite Set Statistics 24
2.3.1	PHD Filter 27
2.3.2	CPHD Filter 31

2.4	Clutter Model	33
2.5	Optimal Subpattern Assignment Metric	34
2.6	Numerical Simulation	35
2.6.1	Test Case Description	36
2.6.2	Results and Discussion	38
2.7	Chapter Summary	43
3	Information Theoretic Sensor Allocation	45
3.1	Background	47
3.1.1	Review of CPHD Filter	47
3.1.2	Development of Information Gain Equations	51
3.1.3	Sensor Tasking Definition	58
3.2	Numerical Simulation	66
3.2.1	Evolution Of Information Gain In Time	66
3.2.2	Test Case Description	68
3.2.3	Simulation Results	70
3.3	Chapter Summary	76
4	Initial Orbit Determination and Follow-On Tracking	77
4.1	Constrained Admissible Region	79
4.2	CPHD Filter Implementation	84
4.2.1	Computation of New Target Probability	84
4.2.2	Measurement-Based Birth Model	86
4.2.3	Augmented CPHD Filter	87
4.2.4	CPHD Clustering	89
4.3	Incorporation of Sensor Allocation Scheme	90
4.3.1	Review of Information Theoretic Sensor Tasking	90
4.4	Numerical Simulation	93

4.4.1	General Test Conditions and Sensor Parameters	93
4.4.2	Test Case 1: Single Thread CPHD	97
4.4.3	Test Case 2: Augmented CPHD with Missed Detections	100
4.4.4	Test Case 3: Augmented CPHD with False Alarms	103
4.4.5	Test Case 4: Augmented CPHD with Ten Objects	106
4.4.6	Limitations of the Method	108
4.5	Chapter Summary	109
5	Conclusion	111
5.1	Research Summary	111
5.2	Future Research Considerations	112
	Bibliography	114
	Appendix	
A	Nomenclature	121
A.1	Single Target Estimation	121
A.2	Multitarget Estimation	122
A.3	Orbits and Measurements	123
A.4	Sensor Allocation	123
A.5	Index Conventions	123
B	Filter Algorithms	124
B.1	AEGIS PHD Filter	124
B.2	AEGIS CPHD Filter	129
C	Information Gain Equations	135
C.1	Derivation of Rényi Divergence	135

C.2	Simplification of Single Target Case	137
C.3	Computation of Numeric Integral	138
C.4	Interpreting Units of Information Gain	140
D	Probabilities for the CPHD Filter Birth Model	143
D.1	Computation of False Alarm Probability	143
D.1.1	Poisson Clutter Model	144
D.1.2	GMM Clutter Model	145
D.1.3	Combined Clutter Model	146
D.2	Computation of New Target Probability	148
E	Spawning Model for the CPHD Filter	149
E.1	Mathematical Background	149
E.2	Derivation of Poisson Spawning Model	151
E.3	Implementation and Results	155

Tables

Table

2.1	Three Component GMM Splitting Library	15
2.2	Sample Hypothesis Table	20
2.3	Station Location in ECEF	36
2.4	Initial State Parameters (EchoStar1)	37
2.5	Initial State and Measurement Standard Deviations	37
2.6	Process Noise Standard Deviations in RIC Frame	37
2.7	Filter Parameters	37
3.1	Sensor Parameters [79]	60
3.2	GEODSS Sensor Parameters [21]	60
3.3	Tasking Modes	66
3.4	Initial State Uncertainties	66
3.5	Initial State Uncertainties	70
3.6	Results Summary	72
4.1	Search-Detect-Track Algorithm	92
4.2	Maui GEODSS Sensor Location and Parameters [79, 21, 35]	94
4.3	Orbital Elements and Longitude at Epoch	95
4.4	Test Case 1 Parameters	97
4.5	Test Case 1 Results	99

4.6 Test Case 2 Results 102

4.7 Test Case 3 Results 105

4.8 Test Case 4 Results 107

Figures

Figure

2.1	Three Component GMM Approximation	15
2.2	MHT Example	20
2.3	Hypothesis Tree	22
2.4	Measurement Visualization	38
2.5	Execution Times	39
2.6	MHT 10 Object OSPA Results	39
2.7	MHT 10 Object Results	40
2.8	JPDA 10 Object Results	41
2.9	PHD 10 Object Results	42
2.10	CPHD 10 Object Results	42
3.1	Illustration of Sensor Field of View	50
3.2	Information Gain Test Case	56
3.3	Information Gain Test Results	57
3.4	Sensor Tasking Grid	58
3.5	Maui Reduced FOR Grid	60
3.6	Sensor Tasking and Filter Block Diagram	62
3.7	Sample Multistep Assignment	63
3.8	Multistep Assignment Windows	64

3.9	Evolution of Information Gain	67
3.10	Object Catalog Characterization	69
3.11	Ground Station Coverage and Object Locations at Epoch	69
3.12	Simulation Position Error Results	71
3.13	Number of Measurements and Objects Detected	72
3.14	Individual Object Errors and Average	73
3.15	Position Error Results of Detected Objects	75
4.1	Constrained Admissible Region	83
4.2	Augmented CPHD Filter	88
4.3	Search-Detect-Track Algorithm	92
4.4	Maui Reduced Box FOR	94
4.5	Single Thread CPHD Test Results	98
4.6	Test ST1C Results	99
4.7	Test ST1D Results	100
4.8	Missed Detection Test OSPA Results	101
4.9	Missed Detection Test Results	101
4.10	Missed Detection $p_D=0.90$ Test Results	103
4.11	Clutter Test OSPA Results	104
4.12	Clutter Test Results	104
4.13	Clutter $\lambda_{\kappa} = 0.01$ Test Results	105
4.14	Ten Object OSPA Results	106
4.15	Number of Objects	106
4.16	Ten Object $C = 5$ OSPA Errors	107
4.17	Ten Object $C = 5$ Results	108
C.1	Numerical Integration Test Case	140
C.2	Rényi Divergence	141

E.1 CPHD Spawning OSPA Results 156

E.2 CPHD Spawning Cardinality Results 156

Chapter 1

Introduction

1.1 Motivation

The past several decades have seen dramatic advancements in space flight, from the first satellite launches, to the Apollo program, to the near ubiquitous use of Earth orbiting satellites for communications, weather, GPS, and other services critical to the functioning of modern society. Along with this increased presence in space has come an increase in the amount of debris left in orbit, due to collisions, planned and unplanned separation events, and retirement of satellites that are no longer functional. This increase in debris poses a serious hazard to the continued use of near Earth space.

Debris objects involved in collisions can damage or destroy active satellites, and repositioning to avoid collisions requires fuel and reduces mission lifetimes. Smaller debris is dealt with by adding shielding to spacecraft instead of performing avoidance maneuvers; however, this increases spacecraft mass. Debris also affects mission planning by requiring an end of life strategy to deorbit or reposition satellites to remove them from the operational environment. The net effect of orbital debris is increased cost and risk to future spacecraft missions.

To mitigate the risks associated with orbital debris, it is essential to accurately determine where objects are and their associated uncertainties. The problem of maintaining an up-to-date catalog of this knowledge and characterizing the debris field for future missions comprises the basic

elements of Space Situational Awareness (SSA).¹ The SSA problem is challenging because there are a limited number of sensors available to gather information on the rapidly growing number of debris objects. The ability to track a large number of objects when only sparse measurement data are available is therefore fundamental to the orbital debris problem. As of 2016, the European Space Agency (ESA) estimates that there are 29,000 objects larger than 10 cm in orbit, with another 670,000 greater than 1 cm.² The National Aeronautics and Space Administration (NASA) orbital debris program office reports that of these, 17,700 are currently tracked by the U.S. Space Surveillance Network (SSN) [54]. The catalog of tracked objects is expected to grow dramatically over the next decade as more advanced sensors are brought online and new launches and breakup events produce more objects to track.

The geosynchronous (GEO) orbit regime is of particular interest because it is heavily used and densely populated by space objects. For satellite slot assignments, the GEO belt is divided into bins of approximately 0.1 degrees in longitude and latitude, and it has become common practice to place more than one satellite in each bin, with typical separation distances between 1 and 100 km [71, 69]. Observations are typically collected from ground-based optical sensors, which are limited to tracking objects greater than one meter in size, though smaller objects have been detected [67]. ESA has cataloged nearly 1500 objects in GEO orbit as of 2016 [22]. New space-based space surveillance (SBSS) sensors such as the Geosynchronous Space Situational Awareness Program (GSSAP) will provide additional and more accurate measurements from a variety of viewing angles, allowing smaller objects to be tracked [39, 55, 7]. While increasing the number and quality of measurements is desirable, it will also increase the number of debris objects in the catalog and add to the complexity of the overall multitarget estimation problem.

¹ This research is concerned with non-cooperative tracking of debris objects, and does not consider tracking cooperative satellites.

² Estimated orbital debris population from http://www.esa.int/Our_Activities/Space_Engineering_Technology/Clean_Space/How_many_space_debris_objects_are_currently_in_orbit, retrieved 28 Sept 2016.

1.2 Challenges in Space Situational Awareness

The main challenges in SSA arise from the fact that there are a large number of objects, a limited number of sensors, and both the dynamical propagation and measurement relationships are nonlinear. The problem is inherently computationally complex, so selection of a filtering scheme must include consideration of computational burden. The limited number of sensors means there are often long gaps between measurements for individual targets. Long gaps increase propagation uncertainty and exacerbate the effects of nonlinear dynamics, which make the assumption of Gaussian state uncertainty less valid [17]. Additionally, the measurement errors themselves may be non-Gaussian. The sensors produce noisy measurements and have limited fields of view, so they must be tasked in a manner to gain as much information as possible on the full multitarget state, while accounting for the need to detect and track newly discovered objects. They can also produce false detections or miss detections of objects that are in the field of view.

Objects in the GEO orbit regime present further challenges, as they are far from ground-based sensors and therefore difficult to observe using radar, precluding the availability of range measurements to this point in time [75]. Optical telescopes provide only angle and angular rate measurements and are further constrained by the lighting conditions on the sensor and objects, i.e., they are generally tasked to observe objects at night and cannot detect objects in Earth's shadow [75]. This limits the length of observation windows, an issue compounded by the slow evolution of GEO orbits relative to ground stations, which necessitates sufficient temporal distribution of measurements to produce accurate estimates. These considerations make the allocation of limited sensor resources more difficult and more important to the fundamental task of tracking geosynchronous space objects.

Finally, the use of angles-only measurements complicates the process of initial orbit determination (IOD) for GEO objects as there is not enough information to compute an orbit solution based on a single observation epoch. Traditional solutions to the problem require measurements from at least three points in time, preferably separated by several hours [79]. In the context of

multitarget filtering the problem becomes more complex, as such measurements have to be stored and compared against subsets of other measurements to see if they yield a viable IOD solution. Together the challenges of numerous and densely spaced objects, limited observation opportunities, and the lack of information to easily initiate new tracks in the filter contribute to the difficulty of solving the multitarget estimation problem for GEO SSA.

1.3 Overview of Research

Multitarget estimation itself is a multifaceted field of research with many applications, including air traffic control, radar, sonar, oceanography, robotics, and image processing [5, 24, 14, 38]. The primary function of multitarget filtering is to resolve measurement ambiguity, determining the correspondence between measurements and objects while accounting for the possibilities of false measurements, or clutter, and missed detections due to imperfect sensors. A number of approaches to the problem exist, including Bayesian and non-Bayesian estimation techniques, implemented in batch, recursive, or fixed-lag processing schemes [62].

The simplest recursive Bayesian multitarget filter is the global nearest neighbor (GNN), or 2D assignment algorithm [4]. In this approach, the single most likely assignment of measurements to targets is used to update individual object states. The assignment is determined by minimizing a global association distance between measurements and targets, accounting for uncertainty. GNN is easily implemented but does not perform well when target or clutter density is high, as is often the case for SSA, due to the increased possibility for incorrect associations.

A more thorough approach to the multitarget estimation problem is Multiple Hypothesis Tracking (MHT), which considers many hypothetical solutions to the 2D assignment problem. The measurement-oriented MHT filter developed by Reid [63] considers all possible measurement-to-track associations, creating a new hypothesis of object state updates for each permutation. Each hypothesis is propagated to the next measurement time, and the process of generating new hypotheses is repeated. If all permutations are kept, one hypothesis will contain the correct associations through all times, and the resulting estimate will be optimal. However, the procedure can make

the problem computationally intractable if the number of objects is large. More recent advances seek to reduce the number of hypotheses [14] or employ track-oriented MHT [41, 6] to address this issue.

While the Reid algorithm considers measurement assignments at a single time, MHT can also be formulated to consider assignments across multiple measurement epochs, or scans. Multi-scan MHT has NP-hard computational complexity, meaning it scales exponentially with the number of targets and measurements. Poore and Robertson have shown that implementation using Lagrangian relaxation can reduce computational burden [60, 61], and in the best case reduces the problem to a series of 2D assignments. Recent work has shown promising results for multi-scan MHT as applied to SSA in terms of accuracy and computational efficiency [1]. In general, many forms of MHT exist and a comparison of all of them is considered outside the scope of this dissertation, therefore the use of multi-scan MHT is not examined in further detail.

Another common single scan approach, Joint Probabilistic Data Association (JPDA), uses all measurements to update all object states, in proportion to the probabilities that individual measurements are associated with each object [24]. By making use of all available information, the method can potentially retain the accuracy of MHT while also being more computationally efficient. However, JPDA filters suffer decreased performance in the presence of densely spaced targets as a result of merging information from measurements other than those truly generated by each target [5, 62].

The methods discussed thus far, GNN, MHT, and JPDA, comprise the basic forms of classical multitarget estimators, though a number of variations have been introduced. The survey paper by Pulford provides an overview of these conventional multitarget filters and their relative performance [62]. More recent developments in multitarget estimation are based in point process theory, from which a simplified mathematical framework known as Finite Set Statistics (FISST) has been derived [47, 49]. FISST is a generalization of single target Bayesian estimation for the multitarget problem, extending familiar concepts such as probability density functions (PDFs) and integral calculus to account for multiple objects. The basic element of FISST is the Random Finite Set (RFS), an order-

independent set of random vectors that is used to define the multitarget state or measurement set at any given time. By defining multitarget PDFs and set integrals, FISST allows Bayes Theorem to be applied directly to the multitarget problem.

From this framework, it is possible to make simplifications by taking moment approximations of the multitarget PDF, similar to the Kalman filter in single target estimation. The first moment approximation is known as the Probability Hypothesis Density (PHD), and produces a filter that is computationally efficient and accurate in simple multitarget tracking scenarios [45]. As a result of truncating information in the multitarget PDF related to target number, the filter is known to have high variability in estimating the number of objects, an issue that is addressed in the Cardinalized PHD (CPHD) filter [48]. The CPHD filter maintains an estimate of the cardinality in addition to the PHD, allowing it to function in environments of densely spaced objects, missed detections, and clutter, all of which are typical of the SSA problem. Both filters require further simplification to be implemented, and the PHD is generally approximated using a Gaussian Mixture (GM) or particle filtering approach [80, 83, 49].

Point process theory and FISST provide a mathematical framework to rigorously address the challenges described above. FISST formalizes concepts from single target statistics for use in multitarget problems, providing a top-down derivation of filters that are computationally efficient and can explicitly account for clutter and missed detections. As explained below and in subsequent chapters, FISST-based filters can also be augmented with initial orbit determination methods to track newly discovered objects and can be used to optimally task sensors to gain information on the multitarget system. The first task is to demonstrate that FISST is a viable approach to the SSA problem by conducting a quantitative comparison against existing multitarget filters.

1.3.1 Comparison of Multitarget Filtering Methods

As discussed, a wide range of options exist to approach multitarget estimation problems. Conventional filters such as MHT and JPDA rely on data association algorithms that can be computationally expensive or error prone [49]. As an alternative, FISST provides a mathematical

framework to derive true multitarget filters from Bayes Theorem, using order-independent sets of random vectors to describe object states and measurement sets. The filters do not require data association algorithms to match measurements to individual objects, which limits their ability to track specific objects of interest, but also produces filters that are more computationally efficient than conventional multitarget methods, and therefore provides an excellent option for SSA [49].

Previous research has examined the application of FISST-based filters to orbit determination and SSA [34, 11], and a limited comparison study has been performed [27]. This dissertation extends this work by providing a quantitative comparison between the FISST-based methods and conventional multitarget filters, including test cases designed to challenge the filters with measurement issues typical of the SSA problem, such as missed detections and clutter. Implementation of the filters is discussed in Ch. 2 and simulation results are provided to demonstrate the relative performance of each.

1.3.2 Information Theoretic Sensor Allocation

Tracking geosynchronous objects from ground-based sensors is a complicated task due to the large number of objects and limited number of available sensors, which are further restricted in the times they are operational. To maintain an accurate catalog, it is essential to collect measurements efficiently, in a manner to maximize use of the limited information available. Recent advances in information theory allow formulation of sensor allocation schemes in terms of information gain functionals computed from the multitarget state and hypothesized measurements [66]. These functionals quantify the difference between the prior and posterior PDF, and therefore account for the reduction in uncertainty for each object as a result of computing a measurement update. Information theoretic tasking schemes favor opportunities to observe multiple objects, as well as objects with large uncertainties in the in-track and cross-track directions where angles-only measurements provide the most benefit.

This dissertation develops a tasking scheme designed to take the fullest advantage of limited observation opportunities by maximizing the information gain. Simulation results are presented

in which a representative catalog of nearly 1000 geosynchronous objects are tracked using three sensors, and demonstrate that the method is able to schedule observations of all objects successfully where other ad hoc methods fall short.

1.3.3 Initial Orbit Determination and Follow-On Tracking

The final subject of research in this dissertation is solution of the search-detect-track problem, in which sensors are tasked to find new objects, which are then initialized in the filter and scheduled for follow-on tracking. The initiation of tracks for newly discovered objects presents unique challenges in FISST-based filters. Classical IOD solutions require measurements for a single object from several different times in order to uniquely determine a six parameter orbit. However, the time history of measurement-to-track associations is not generally available in FISST filters, so this is not possible. Formulations exist to include track labeling in FISST [57], but to produce an IOD solution the filters would need to solve a multi-scan assignment problem, which is not considered in this study. Instead, the approach is to instantiate a set of potential initial orbit solutions based on a measurement set at one point in time and use follow-on tracking to refine the estimate.

The proposed method is based on approximating a constrained admissible region (CAR) of possible IOD solutions at the time when the new object is detected [19]. The filter is initialized with a weighted mixture of these solutions, which are refined by future measurements. The previously developed information theoretic tasking scheme is augmented by a threshold which allows it to switch from searching for new objects to tracking known objects in the filter. The uncertainties associated with newly initiated tracks automatically drive follow-on tracking due to the expected information gain.

1.3.4 Contributions of Research

The primary contributions of the dissertation are therefore summarized as follows:

1. Finite Set Statistics is demonstrated as a viable framework for developing multitarget filters for GEO SSA. Filters derived from FISST, in particular the CPHD filter, achieve

comparable performance as compared to conventional multitarget filtering methods.

2. An information theoretic sensor allocation scheme is developed for use with the CPHD filter. The scheme achieves superior performance relative to less rigorous approaches.
3. A solution to the search-detect-track problem is presented, leveraging the use of the constrained admissible region to initiate tracks and the previously developed information theoretic sensor tasking scheme to schedule follow-on tracking.

1.4 Thesis Overview

This dissertation is organized as follows. Chapter 2 discusses the implementation of four multitarget filters, MHT, JPDA, PHD, and CPHD, and examines the relative performance of each in a simulated SSA tracking scenario. The chapter begins with a review of single target filtering, including an overview of the AEGIS Gaussian Mixtures scheme that forms the backbone of the various multitarget filters [18]. A brief description of Finite Set Statistics is provided prior to descriptions of the PHD and CPHD filters. Complete algorithms for the AEGIS PHD and CPHD filters are included in Appendix B.

Chapter 3 describes the use of information theoretic sensor tasking for GEO SSA and provides simulation results demonstrating the benefits of the scheme as compared to simpler ad hoc methods. Two different information gain functionals are considered in a simplistic scenario, with the Rényi divergence selected as more appropriate for the problem under consideration. The use of single-step and multistep sensor tasking schemes are described, along with simple grid-based search approaches. Derivations of relevant information gain formulas and specific implementation details are included in Appendix C.

Chapter 4 demonstrates a solution to the search-detect-track problem in which tracks are initiated for newly discovered objects and follow-on tracking is automatically scheduled using the information theoretic tasking scheme. The principles of the constrained admissible region IOD solution are provided, and different approaches to incorporating new targets in the filter are con-

sidered in a simple tracking scenario. Simulations are included to demonstrate the viability of the solution. Derivation and demonstration of an alternate IOD approach that models targets spawned from known objects is presented in Appendix E.

Chapter 5 provides a summary of results and conclusions, as well as directions for future research.

Chapter 2

Comparison of Multitarget Filtering Methods

This chapter provides a review of multitarget filtering techniques and a quantitative comparison between the methods as applied to SSA. Multitarget estimation seeks to determine the most likely number of objects and their states, based upon given measurements. The assumed equations of motion for each object are

$$\dot{\boldsymbol{x}} = F(\boldsymbol{x}, \boldsymbol{v}, t) \quad (2.1)$$

where \boldsymbol{x} is the state vector and \boldsymbol{v} is a zero mean white noise process. The measurement relationship is given by

$$\boldsymbol{z}_k = G(\boldsymbol{x}_k, t_k) + \boldsymbol{\epsilon}_k \quad (2.2)$$

where $\boldsymbol{\epsilon}_k$ is a zero mean measurement noise vector, and the convention $\boldsymbol{x}_k = \boldsymbol{x}(t_k)$ has been adopted. For the orbital debris problem, both the dynamics and measurement relationships are nonlinear in general.

The remainder of this chapter provides details on the implementation of several multitarget filters, an explanation of metrics used to evaluate their performance, and results from a simulation comparing the filters. In developing and implementing the filters, several simplifying assumptions are made and held throughout this chapter. The number of objects in all cases is considered fixed, no target birth or death is considered, though subsequent research presented in Ch. 4 does examine estimation of new targets. Each target is assumed to produce at most one measurement at a given time, and each measurement represents at most one target. To establish a common notation and

framework for discussing multitarget estimation, a review of single target filtering is first included, following the development in Mahler [49].

2.1 Review of Single Target Filtering

The most general formulation of the single target filtering problem is based on Bayes Theorem, which provides a means to update the PDF of an object's state, given the PDFs of the predicted *a priori* state and measurements. Full Bayesian estimation is not generally practical; however, it provides the basis to make simplifications, resulting in more easily implemented estimators such as the Kalman filter. The development and nomenclature provided in this section will be used as a starting point for discussion of conventional multitarget methods, which are based on using multiple single target filters, as well as to draw parallels in the development of the FISST-based filters discussed later.

The Bayes filter begins with the PDF, $p_{k-1}(\mathbf{x}_{k-1}|z_{1:k-1})$, describing the likelihood that the target has state \mathbf{x}_{k-1} at time t_{k-1} , conditioned on the time series of measurements $z_{1:k-1} = z_1, \dots, z_{k-1}$. The first task is to define the Markov transition density, which describes the evolution of the PDF due to system dynamics and is denoted $f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1})$. The *a priori* PDF is computed by taking the product of the Markov transition density and prior PDF and integrating over all possible prior states.

$$p_{k|k-1}(\mathbf{x}_k|z_{1:k-1}) = \int f_{k|k-1}(\mathbf{x}_k|\mathbf{x})p_{k-1}(\mathbf{x}|z_{1:k-1})d\mathbf{x} \quad (2.3)$$

Next, the likelihood density function $g_k(\mathbf{z}_k|\mathbf{x}_k)$ is needed, which describes the probability that the measurement \mathbf{z}_k is generated by state \mathbf{x}_k . The product $g_k(\mathbf{z}_k|\mathbf{x}_k) \cdot p_{k|k-1}(\mathbf{x}_k|z_{1:k-1})$ conditions the likelihood based on the *a priori* PDF. It must be normalized to produce the correct posterior PDF, which results in the following realization of Bayes Theorem.

$$p_k(\mathbf{x}_k|z_{1:k}) = \frac{g_k(\mathbf{z}_k|\mathbf{x}_k)p_{k|k-1}(\mathbf{x}_k|z_{1:k-1})}{\int g_k(\mathbf{z}_k|\mathbf{x})p_{k|k-1}(\mathbf{x}|z_{1:k-1})d\mathbf{x}} \quad (2.4)$$

The Kalman filter exists in many forms, but is most generally a second moment approximation to the Bayes filter [31], which means it can be derived by assuming that the first and second moments

(mean and covariance) sufficiently capture the information of the full Bayes posterior PDF.

$$p_k(\mathbf{x}_k|z_{1:k}) \approx p_k(\mathbf{x}_k; \hat{\mathbf{x}}_k, P_k) \quad (2.5)$$

The Kalman filter therefore propagates and corrects the estimated mean state $\hat{\mathbf{x}}_k$ and covariance P_k defined by

$$\hat{\mathbf{x}}_k = \int \mathbf{x}_k \cdot p_k(\mathbf{x}_k|z_{1:k}) d\mathbf{x}_k \quad (2.6)$$

$$P_k = \int (\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T \cdot p_k(\mathbf{x}_k|z_{1:k}) d\mathbf{x}_k \quad (2.7)$$

The Kalman filter is a suitable option for linear dynamical systems because an exact closed-form solution is available for mapping the mean and covariance under linear transformations. Variations of the Kalman filter have been derived to address nonlinear problems, including the Extended Kalman Filter (EKF) [28] and the Unscented Kalman Filter (UKF) [36, 37]. However, both filters assume that state and measurement errors are Gaussian-distributed, which is not generally the case in orbit determination. The rest of this section describes the implementation of a modified UKF, the Adaptive Entropy-based Gaussian-mixture Information Synthesis (AEGIS) filter, which uses multiple Gaussian distributions to represent the uncertainty of a single object [17, 18]. The filter is an appropriate choice for SSA problems because it is nonlinear and it allows for non-Gaussian uncertainty propagation during long data gaps, which are common in SSA.

2.1.1 AEGIS UKF

The Kalman filter is derived by assuming state and measurement errors are well described by the Gaussian, or normal, distribution. This assumption is justified by the central limit theorem [74], which states that the distribution of the mean error of a large number of samples will tend to be distributed normally. In addition, it produces a closed-form solution to predict and update the state PDF in linear systems because Gaussian PDFs are maintained under linear transformations.

In a system with nonlinear dynamics, such as that encountered in orbit determination, the evolution of a Gaussian PDF in time is non-Gaussian, and the prediction of the assumed distribution departs from the true state uncertainty over long time intervals. This problem can be addressed

by approximating the state PDF as a linear combination of Gaussian PDFs to produce a Gaussian Mixture Model (GMM), which can be structured to converge to an arbitrary PDF by using an increasing number of Gaussian components [72]. The GMM components are then propagated and updated using the Kalman filter equations. The following development of the AEGIS UKF follows that given by DeMars [17].

The Gaussian distribution of a random vector \mathbf{x} with mean $\hat{\mathbf{x}}$ and covariance P is defined as

$$p_g(\mathbf{x}; \hat{\mathbf{x}}, P) = \frac{1}{\sqrt{|2\pi P|}} \exp \left[-\frac{1}{2}(\mathbf{x} - \hat{\mathbf{x}})^T P^{-1}(\mathbf{x} - \hat{\mathbf{x}}) \right] \quad (2.8)$$

The GMM approximation of an arbitrary PDF can be computed as a weighted sum of Gaussian PDFs

$$p(\mathbf{x}) \approx \sum_{j=1}^J w_j p_g(\mathbf{x}; \hat{\mathbf{x}}_j, P_j) \quad (2.9)$$

where the weights w_j must adhere to the properties $w_j \geq 0$ and $\sum_{j=1}^J w_j = 1$.

In order to apply the GMM approach in a filter, a test is needed to determine when the predicted state PDF is becoming non-Gaussian. This test can be implemented using the differential entropy, which for a Gaussian distribution is defined as

$$H(\mathbf{x}) = \frac{1}{2} \log |2\pi e P| \quad (2.10)$$

The linear time rate of change of the differential entropy is computed from

$$\dot{H}(\mathbf{x}) = \text{trace}\{A(\mathbf{x}(t), t)\} \quad (2.11)$$

$$A(\mathbf{x}(t), t) = \left[\frac{\partial F}{\partial \mathbf{x}} \right]_{\mathbf{x}=\hat{\mathbf{x}}} \quad (2.12)$$

The linearly propagated entropy can be compared to the entropy computed from the nonlinear UKF propagation of the PDF. If the difference is above a threshold, it indicates that the state uncertainty is becoming non-Gaussian. At this point, the initial Gaussian distribution can be split into several components, the time evolution of which will produce a GMM PDF that is non-Gaussian and better matches the state uncertainty.

The procedure to split a single Gaussian PDF into several PDFs is implemented by defining a splitting library for the univariate standard normal distribution, which can then be generalized to all Gaussian distributions. The library provides the weights, means, and standard deviations for a GMM approximation with J components, and is generated by solving a constrained optimization problem to most accurately match the original PDF. The library for a 3-component GMM is provided in Table 2.1 [17], and a simple illustration of the technique is provided in Figure 2.1, in which three components are used to approximate the standard normal distribution.

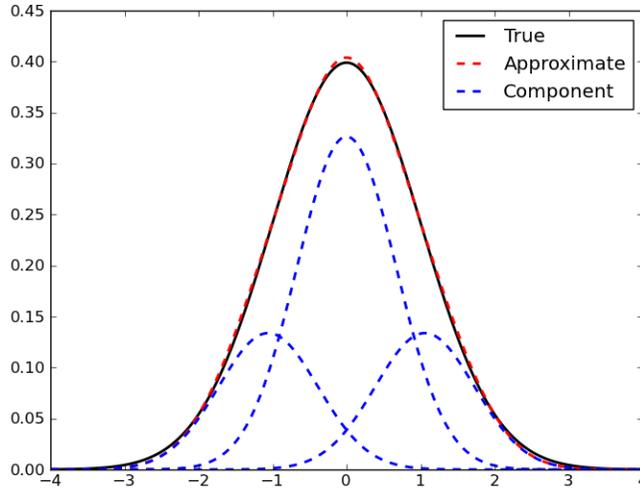


Figure 2.1: Three Component GMM Approximation

Table 2.1: Three Component GMM Splitting Library

i	$\tilde{\alpha}_j$	\tilde{m}_j	$\tilde{\sigma}_j$
1	0.2252246249136750	-1.057515461475881	0.6715662886640760
2	0.5495507501726501	0	0.6715662886640760
3	0.2252246249136750	1.057515461475881	0.6715662886640760

To create a multivariate GMM approximation from the univariate splitting library, the spectral factorization of the covariance is needed. The multivariate distribution can be split along a single axis using the following equations. Assume the current Gaussian component has weight,

mean, and covariance given by w , $\hat{\mathbf{x}}$, and P . The spectral decomposition is given by

$$P = V\Lambda V^T \quad (2.13)$$

where Λ is the diagonal matrix of eigenvalues and the columns of V are the corresponding eigenvectors. Splitting along the k -th eigenvector produces the following updated weights, means, and covariances.

$$w_j = \tilde{\alpha}_j w \quad (2.14)$$

$$\hat{\mathbf{x}}_j = \hat{\mathbf{x}} + \sqrt{\lambda_k} \tilde{m}_j \mathbf{v}_k \quad (2.15)$$

$$P_j = V\Lambda_j V^T \quad (2.16)$$

$$\Lambda_j = \text{diag}\{\lambda_1, \dots, \tilde{\sigma}_j^2 \lambda_k, \dots, \lambda_L\} \quad (2.17)$$

where λ_k and \mathbf{v}_k are the k -th eigenvalue and eigenvector, and L is the number of dimensions. The distribution can be split along multiple dimensions by repeating the procedure above. For simplicity, this research implements splitting in a single dimension, along the eigenvector corresponding to the largest eigenvalue, which represents the direction of the greatest uncertainty.

The filter uses the UKF predictor equations, which require definition of the sigma point matrix and weights associated with the unscented transform. The weights are given by [85]

$$W_0^m = \lambda / (L + \lambda) \quad (2.18)$$

$$W_0^c = \lambda / (L + \lambda) + (1 - \alpha^2 + \beta) \quad (2.19)$$

$$W_l^m = W_l^c = 1 / \{2(L + \lambda)\} \quad l = 1, \dots, 2L \quad (2.20)$$

with associated tuning parameters $10^{-4} \leq \alpha \leq 1$, $\beta = 2$, $\kappa = 3 - L$, and $\lambda = \alpha^2(L + \kappa) - L$. The sigma point matrix is generated from the component's mean and the square root of the covariance, computed using the Cholesky decomposition such that $P = (\sqrt{P})(\sqrt{P})^T$.

$$\chi_{k-1} = \left[\hat{\mathbf{x}}_{k-1} \quad \hat{\mathbf{x}}_{k-1} + \gamma \sqrt{P_{k-1}} \quad \hat{\mathbf{x}}_{k-1} - \gamma \sqrt{P_{k-1}} \right]_{L \times (2L+1)} \quad (2.21)$$

where $\gamma = \sqrt{(L + \lambda)}$. The *a priori* mean and covariance are computed by propagating the sigma points according to the system dynamics.

$$\bar{\chi}_{k,l} = \phi(\chi_{k-1,l}, \mathbf{v}_{k-1}, t_{k-1}) \quad (2.22)$$

$$\bar{\mathbf{x}}_k = \sum_{l=0}^{2L} W_l^m \bar{\chi}_{k,l} \quad (2.23)$$

$$\bar{P}_k = Q_k + \sum_{l=0}^{2L} W_l^c (\bar{\chi}_{k,l} - \bar{\mathbf{x}}_k)(\bar{\chi}_{k,l} - \bar{\mathbf{x}}_k)^T \quad (2.24)$$

where $\phi(\cdot)$ is the solution of the dynamics Eq. (2.1) and $\bar{\chi}_{k,l}$ is the l -th column of the sigma point matrix. After incorporating the process noise covariance, the sigma points are recomputed from the predicted mean and covariance.

$$\bar{\chi}_k = \begin{bmatrix} \bar{\mathbf{x}}_k & \bar{\mathbf{x}}_k + \gamma\sqrt{\bar{P}_k} & \bar{\mathbf{x}}_k - \gamma\sqrt{\bar{P}_k} \end{bmatrix}_{L \times (2L+1)} \quad (2.25)$$

If at any point during the propagation nonlinearity is detected in a GMM component, it is split according to the procedure above.

The measurement update equations are similarly based on the UKF, with a separate update computed for each GMM component. The measurement sigma points and expected measurement vector for component j are given by

$$\zeta_k^{(j)} = G(\bar{\chi}_k^{(j)}, t_k) \quad (2.26)$$

$$\bar{\mathbf{z}}_k^{(j)} = \sum_{l=0}^{2L} W_l^m \zeta_{k,l}^{(j)} \quad (2.27)$$

The component mean and covariance updates are computed from the following equations.

$$P_{zz}^{(j)} = R_k + \sum_{l=0}^{2L} W_l^c (\zeta_{k,l}^{(j)} - \bar{\mathbf{z}}_k^{(j)})(\zeta_{k,l}^{(j)} - \bar{\mathbf{z}}_k^{(j)})^T \quad (2.28)$$

$$P_{xz}^{(j)} = \sum_{l=0}^{2L} W_l^c (\bar{\chi}_{k,l}^{(j)} - \bar{\mathbf{x}}_k^{(j)})(\zeta_{k,l}^{(j)} - \bar{\mathbf{z}}_k^{(j)})^T \quad (2.29)$$

$$K_k^{(j)} = P_{xz}^{(j)} [P_{zz}^{(j)}]^{-1} \quad (2.30)$$

$$\hat{\mathbf{x}}_k^{(j)} = \bar{\mathbf{x}}_k^{(j)} + K_k^{(j)} (\mathbf{z}_k - \bar{\mathbf{z}}_k^{(j)}) \quad (2.31)$$

$$P_k^{(j)} = \hat{P}_k^{(j)} - K_k^{(j)} P_{zz}^{(j)} [K_k^{(j)}]^T \quad (2.32)$$

The final step is computation of the updated weights for the GMM. If the *a priori* component weight is given by $\bar{w}_k^{(j)}$, the updated weight is given by

$$g_k^{(j)} = p_g(\mathbf{z}_k; \bar{\mathbf{z}}_k^{(j)}, P_{zz}^{(j)}) \quad (2.33)$$

$$w_k^{(j)} = \frac{g_k^{(j)} \bar{w}_k^{(j)}}{\sum_{l=1}^{J_k} g_k^{(l)} \bar{w}_k^{(l)}} \quad (2.34)$$

After the weights have been updated, the number of components can be reduced by removing components with low weights or merging components that are close together. Merging GMM components is also an important consideration in multitarget filtering, details of this procedure are provided in Section 2.3.1.

2.2 Conventional Multitarget Filtering

Conventional multitarget filtering employs a bottom-up approach, in which the problem is divided into many single target tasks, which can be solved using a filter such as the AEGIS UKF. In order to perform the measurement update for each target, it is necessary to associate individual measurements to targets. For a given set of objects and measurements, multiple measurement-to-track associations can be hypothesized and then weighted using an association distance metric to set the probabilities for each hypothesis.

Having formed the hypotheses, two of the most common association algorithms are Multiple Hypothesis Tracking (MHT) [63] and Joint Probabilistic Data Association (JPDA) [24]. As suggested by the name, MHT maintains a list of hypothetical associations along with an estimate of their probabilities. For a given hypothesis, each object is updated using the appropriate measurement and its state is then propagated to the next time. If all hypotheses are kept, the correct association is guaranteed to be included and the associated state estimates will be optimal. However, as the number of objects increases, the list of hypotheses grows to the point that the problem becomes computationally intractable, therefore techniques are required to avoid generating unlikely hypotheses and remove those that decrease in probability over time. JPDA offers an alternative approach by merging all hypotheses to perform the measurement update. Object states

and uncertainties are updated using weighted residuals conditioned according to the hypotheses' probabilities. In this manner, only one set of object states is propagated for each time step. However, the composite update does include incorrect hypotheses, and for densely spaced objects, state estimate accuracy may be reduced [5].

The prediction step is the same for both filters, and within the AEGIS framework, each GMM component is propagated using Eqs. (2.22)-(2.24), splitting into more components if needed. To perform the measurement update, the data association problem must be solved, as discussed in the following sections.

2.2.1 Multiple Hypothesis Tracking

A full MHT filter employs all permutations of measurement-to-track association, which guarantees the correct association is included, but can become computationally burdensome. The method used in this dissertation is measurement-oriented MHT as originally developed by Reid [63], augmented by N -scan pruning and the Murty algorithm for hypothesis generation in order to manage the number of hypotheses [14, 5, 53]. MHT was originally developed considering only one Gaussian PDF per object, but the implementation remains the same for AEGIS, with associations determined for each measurement-to-component.

The main function of the Reid algorithm is the generation of a table of hypotheses and computation of their associated probabilities. Each column of the table represents possible assignments for a single component. Each row of the table represents one hypothetical assignment of measurements to all components. To illustrate the computation of the hypothesis table and associated probabilities, a simple example is provided for the case of two measurements and two components.

The association distance between a measurement $\mathbf{z}_k^{(i)}$, and predicted component state $\bar{\mathbf{x}}_k^{(j)}$ is computed using the predicted measurement $\bar{\mathbf{z}}_k^{(j)}$ and uncertainty $P_{zz}^{(j)}$.

$$d_{ij}^2 = [\mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^{(j)}]^T [P_{zz}^{(j)}]^{-1} [\mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^{(j)}] \quad (2.35)$$

The association distance is used both to compute the probability of an association, and also to set

gates to preclude unlikely associations from being considered. In this manner, if a measurement is far from an object, it need not be included in the hypothesis table. For this example, assume that both measurements fall within the gate for both components, i.e., all assignments are feasible. The hypothesis table for this example is given in Table 2.2 and a simple illustration provided in Figure 2.2. The estimated object components C1 and C2 and corresponding gates are mapped to the measurement space to allow for comparison to the observations M1 and M2.

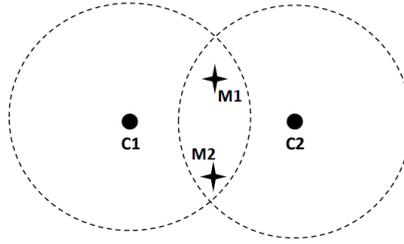


Figure 2.2: MHT Example

Table 2.2: Sample Hypothesis Table

Hyp	C1	C2	$L(\text{Hyp})$
1	0	0	$(1 - p_D)^2 \lambda^2$
2	1	0	$g_{11} p_D (1 - p_D) \lambda$
3	2	0	$g_{21} p_D (1 - p_D) \lambda$
4	0	1	$g_{12} p_D (1 - p_D) \lambda$
5	2	1	$g_{21} g_{12} p_D^2$
6	0	2	$g_{22} p_D (1 - p_D) \lambda$
7	1	2	$g_{11} g_{22} p_D^2$

The indices are defined such that 0 indicates no assignment, and 1 and 2 indicate assignment of the appropriate measurement to the component. In this manner, the first row of the table, [0 0], represents the hypothesis that both measurements are false alarms and are not assigned to either component. The row [2 1] assigns measurement 2 to component 1 and measurement 1 to component 2.

To compute the probability that a given hypothesis is correct, the likelihood of each assign-

ment is needed. The likelihood of a non-assignment is a function of the number of detections within the gate, N_D , the probability of detection, p_D , and the false alarm density, λ . The likelihood for an assignment is additionally a function of the measurement likelihood, g_{ij} .

The likelihood results can be summarized in Eqs. (2.36)-(2.37), where i is the assignment index from the hypothesis table. To compute the probability for each hypothesis, the likelihoods of assignments in a row must be multiplied together. The likelihoods for each row in the hypothesis table are then normalized to produce valid probabilities. Note that computation of the likelihoods results in common factors of λ which have been canceled in Table 2.2 [5].

$$L(i) = \begin{cases} (1 - p_D)\lambda^{N_D} & \text{if } i = 0 \\ g_{ij}p_D\lambda^{(N_D-1)} & \text{if } i \geq 1 \end{cases} \quad (2.36)$$

$$g_{ij} = p_g(\mathbf{z}^{(i)}; \bar{\mathbf{z}}^{(j)}, P_{zz}^{(j)}) \quad (2.37)$$

Having assigned measurements, components can be updated using the AEGIS UKF Eqs. (2.31)-(2.32). In the case of a non-assignment, the *a priori* values of the state and covariance are retained. For each hypothesis, the tracks are updated and propagated to the next epoch. At the new time, a given hypothesis will spawn a new hypothesis table and the process to compute probabilities repeats. Adhering to Bayes Theorem, the probabilities in the new hypothesis table must be multiplied by the probability of the original hypothesis that spawned them in order to keep the full list of estimated probabilities correctly normalized.

In practice, as time progresses, one hypothesis should align better and better with the measurements provided, and eventually emerge as having the highest probability. As mentioned earlier, to keep the number of hypotheses from growing unreasonably large, techniques may be employed to remove unlikely hypotheses and limit the number of new hypotheses spawned for each measurement update. Hypotheses with probabilities below a certain threshold can be removed, as well as those that do not produce additional hypotheses that pass the gate checks. This dissertation also employs N -scan pruning and the Murty algorithm to achieve better computational efficiency [14, 5, 53].

2.2.1.1 N-Scan Pruning

The purpose of N -scan pruning is to allow the filter to progress several steps before making decisions regarding which hypotheses to keep [5]. This allows the assignment of additional measurements to either confirm or disprove hypotheses before removal. The easiest way to visualize the approach is through the creation of a hypothesis tree, as shown in Figure 2.3.

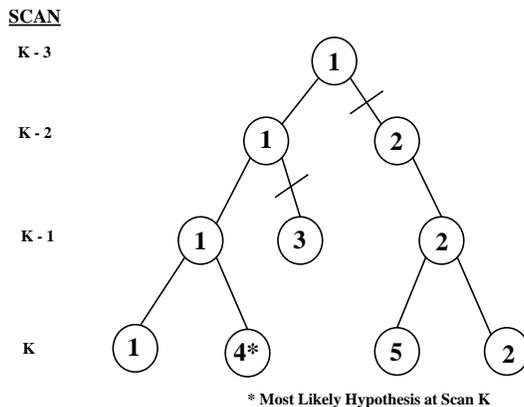


Figure 2.3: Hypothesis Tree

The branches in the tree represent the formation of new hypotheses. After scan $K-1$, hypothesis 3 does not generate any new hypotheses and can therefore be removed. At scan K , there are 4 remaining hypotheses. N -scan pruning is implemented by finding the most likely hypothesis at the current time, then going back N scans and removing any hypotheses that do not share a common branch at that point. In this example, hypothesis 4 is the most likely at time K . By going back 2 scans, it is possible to remove all hypotheses that do not share a common branch with hypothesis 4, in this case 2 and 5. In this manner, the number of hypotheses can be reduced by regularly removing branches that lead to less likely solutions. An alternate implementation uses the sum of all hypotheses from each branch to determine which to remove [14]. In this case it would sum the probabilities of hypotheses 1 and 4 and compare against the sum for hypotheses 2 and 5.

2.2.1.2 Murty Algorithm

The Murty algorithm produces the k -best hypotheses for a given set of measurements and tracks, rather than generating all possible hypotheses that pass the gate checks. This limits the number of new hypotheses that are considered at any given time. The algorithm works by setting up the measurement-to-component association as a 2D assignment problem. The single most likely association can be determined from established methods, such as the auction algorithm [5]. Following this, individual assignments are removed one at a time from the table and the auction algorithm is run again to produce the most likely association from the reduced problem. The process is repeated until a set number of hypotheses have been returned, details of the implementation are available in References [14] and [5].

2.2.2 Joint Probabilistic Data Association

An alternative approach to keep the multitarget problem computationally tractable is Joint Probabilistic Data Association [24]. This method produces the same hypothesis table and probabilities as MHT, but then merges hypotheses to perform a weighted measurement update, meaning only a single set of component means and covariances are propagated to the next time step.

Having computed the probabilities of each hypothesis, the sum of probabilities of all hypotheses matching a given measurement to a component are used to compute the weighted innovation. Returning to the example above, both hypotheses 2 and 7 assign measurement 1 to component 1. The total probability of this assignment is therefore $p_{11} = p(H_2) + p(H_7)$. Denoting the innovation $\mathbf{y}_k^{(i,j)} = \mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^{(j)}$, the weighted innovation accounts for the contributions from all measurements m that may be assigned to the j -th component.

$$\mathbf{y}_k^{(j)} = \sum_{i=1}^m p_{ij} \mathbf{y}_k^{(i,j)} \quad (2.38)$$

The state can then be updated using

$$\hat{\mathbf{x}}_k^{(j)} = \bar{\mathbf{x}}_k^{(j)} + K_k^{(j)} \mathbf{y}_k^{(j)} \quad (2.39)$$

where the Kalman gain $K_k^{(j)}$ is computed from Eq. (2.30), and for simplicity it is assumed R_k is the same for all measurements so K_k is only a function of the predicted component state. The covariance update consists of two terms designed to account for the standard Kalman filter update and to increase the covariance as a result of uncertainty in the association process.

$$P_k^{(j)} = P_{k,0}^{(j)} + dP_k^{(j)} \quad (2.40)$$

The first term accounts for the normal Kalman filter update, and includes the term p_{0j} denoting the probability of a missed detection. In the case a detection is certainly missed, the *a priori* covariance is retained, while in the case a detection is certainly made, the standard covariance update from Eq. (2.32) is used, denoted by $P_k^{*(j)}$.

$$P_{k,0}^{(j)} = p_{0j} \bar{P}_k^{(j)} + (1 - p_{0j}) P_k^{*(j)} \quad (2.41)$$

The second term inflates the covariance as a result of the uncertainty in data association. It is computed from

$$dP_k^{(j)} = K_k^{(j)} \left[\sum_{i=1}^{m_k} p_{ij} \mathbf{y}_k^{(i,j)} [\mathbf{y}_k^{(i,j)}]^T - \mathbf{y}_k^{(j)} [\mathbf{y}_k^{(j)}]^T \right] [K_k^{(j)}]^T \quad (2.42)$$

As a result of this merging effect, JPDA may be considered an extension to MHT, in which hypothesis management is accomplished by reducing the hypothesis table to a single set of update equations. JPDA may also be used in conjunction with the Murty algorithm to limit the number of hypotheses considered and improve computational efficiency. However, in regions of densely spaced targets, JPDA may produce less accurate state estimates than MHT as a result of merging information from incorrect hypotheses in the measurement update [62].

2.3 Finite Set Statistics

The drawback to conventional multitarget filters is their reliance on single target filters to propagate and update object states. While single target filters are well understood, they can only be used to process the measurement that is generated by their object, they have no ability to process a set of measurements from multiple objects at once. This imposes a strict requirement to

associate individual measurements to specific targets, which means the methods amount to a data association algorithm with a bank of Kalman filters running underneath. Data association is the most computationally expensive step of these methods, and can introduce errors if the association is incorrect.

To address these issues, Finite Set Statistics seeks to generalize concepts from single target filtering to multitarget problems, allowing for the development of true multitarget filters. The goal is to apply multitarget statistics to process a set of measurements and update a set of object states free from the data association problem. The full development of FISST and careful definition of multitarget statistics can be found in Reference [49]. For this research, it suffices to say that there exists a multitarget analog for many of the concepts and functions familiar in single target statistics, the most important of which are summarized below.

The basic element of FISST is the Random Finite Set (RFS), which is an order-independent set of random vectors. In the orbit determination problem, RFSs are used to represent sets of object state vectors or measurements. The RFS allows for varying numbers of random vectors and includes the null set to account for the possibility of no targets or measurements. Sample state and measurement RFSs can be defined by

$$X = \{\mathbf{x}_1, \dots, \mathbf{x}_{n_k}\} \quad (2.43)$$

$$Z = \{\mathbf{z}_1, \dots, \mathbf{z}_{m_k}\} \quad (2.44)$$

To illustrate the core concepts of FISST, the rest of this section will provide examples for the state RFS.

The statistical information about the state RFS is contained in a multitarget PDF, which is easiest to define in conjunction with the belief-mass function and set integral. Begin by assuming that the function $p(X) \geq 0$ exists and represents the multitarget PDF. The belief-mass function is the probability that X will be in some space S , and is therefore analogous to the cumulative distribution function used in single target statistics. This function is defined as

$$\beta_X(S) = \Pr(X \in S) = \int_S p(X) dX \quad (2.45)$$

The integral in the above equation requires careful definition, because the RFS contains a discrete number of continuous state vectors. The set integral is therefore defined as a sum of integrals, each term accounting for a different possible number of targets. Over a region S , the set integral is defined as

$$\begin{aligned} \int_S p(X) dX &= \sum_{n=0}^{\infty} \frac{1}{n!} \int_{S \times \dots \times S} j^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) d\mathbf{x}_1 \cdots d\mathbf{x}_n \\ &= p(\emptyset) + \int_S j^{(1)}(\{\mathbf{x}\}) d\mathbf{x} + \frac{1}{2} \int_{S \times S} j^{(2)}(\{\mathbf{x}_1, \mathbf{x}_2\}) d\mathbf{x}_1 d\mathbf{x}_2 + \dots \end{aligned} \quad (2.46)$$

where $j^{(n)}(\cdot)$ is the n -th order Janossy density, representing the spatial distribution of n targets. The multitarget PDF can now be defined as a function whose set integral is 1 when integrated over all states and possible number of targets.

$$\int p(X) dX = 1 \quad (2.47)$$

From the above definition of the set integral, it is useful to define the cardinality distribution $p(n)$, which is the probability that there are exactly n targets in X , or that the cardinality $|X| = n$.

$$p(n) = \Pr(|X| = n) = \frac{1}{n!} \int j^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) d\mathbf{x}_1 \cdots d\mathbf{x}_n \quad (2.48)$$

Therefore, each term in the set integral is the cardinality distribution for that number of targets, and the integral of the PDF can be rewritten as the sum of these probabilities.

$$\int p(X) dX = p(\emptyset) + p(1) + p(2) + \dots = 1 \quad (2.49)$$

The final concept which needs to be extended to the multitarget problem is that of statistical moments. In single target statistics, the first moment of a PDF is the mean, defined by

$$\mathbf{m} = \int \mathbf{x} \cdot p(\mathbf{x}) d\mathbf{x} \quad (2.50)$$

For the multitarget case, the first moment is the Probability Hypothesis Density (PHD) function [45], similarly defined by

$$\nu(\mathbf{x}) = \int p(\{\mathbf{x}\} \cup X) dX \quad (2.51)$$

The PHD represents target density in some region of space, making it possible to estimate the number of targets in the space by integration.

$$\hat{N} = \int_S \nu(\mathbf{x}) d\mathbf{x} \quad (2.52)$$

Note that the PHD is defined in the single target state space instead of the full multitarget space, making it an intuitively easier function to work with than the multitarget PDF.

2.3.1 PHD Filter

Having defined the concepts essential to FISST, this section provides the equations for the simplest FISST-derived estimator, the PHD filter. The PHD filter is a moment approximation to the multitarget Bayes filter, similar to how the Kalman filter is a moment approximation to the single target Bayes filter. It reduces the full multitarget PDF at each step to the simpler PHD function. The process is summarized as a two step predictor/corrector, depicted below [49].

$$\begin{array}{ccccccc}
 & & & \text{predictor} & & \text{corrector} & \\
 \text{Bayes:} & \dots & \rightarrow & p_{k-1}(X_{k-1}|Z_{1:k-1}) & \rightarrow & p_{k|k-1}(X_k|Z_{1:k-1}) & \rightarrow & p_k(X_k|Z_{1:k}) & \rightarrow & \dots \\
 & & & \downarrow & & \downarrow & & \downarrow & & \\
 \text{PHD:} & \dots & \rightarrow & \nu_{k-1}(\mathbf{x}_{k-1}) & \rightarrow & \nu_{k|k-1}(\mathbf{x}_k) & \rightarrow & \nu_k(\mathbf{x}_k) & \rightarrow & \dots
 \end{array}$$

The full PHD recursion is given by Eqs. (2.53)-(2.54), and as before, does not include terms for target birth which will be considered in Ch. 4. The corrector equation here is for the single sensor case, and can be extended to include multiple sensors.

$$\nu_{k|k-1}(\mathbf{x}) = \int p_S(\mathbf{x}) f_{k|k-1}(\mathbf{x}_k|\mathbf{x}) \nu_{k-1}(\mathbf{x}) d\mathbf{x} \quad (2.53)$$

$$\begin{aligned}
 \nu_k(\mathbf{x}) = & [1 - p_D(\mathbf{x}_k)] \nu_{k|k-1}(\mathbf{x}) \\
 & + \sum_{\mathbf{z}_k \in Z_k} \frac{p_D(\mathbf{x}_k) g_k(\mathbf{z}_k|\mathbf{x}_k) \nu_{k|k-1}(\mathbf{x})}{\kappa_k(\mathbf{z}_k) + \int p_D(\mathbf{x}) g_k(\mathbf{z}_k|\mathbf{x}) \nu_{k|k-1}(\mathbf{x}) d\mathbf{x}}
 \end{aligned} \quad (2.54)$$

where $p_D(\mathbf{x}_k)$ is the probability of detection, $\kappa_k(\mathbf{z}_k)$ is the PHD of false alarms, and $p_S(\mathbf{x})$ is the probability of target survival from time t_{k-1} to t_k , which accounts for all persisting targets and implicitly the possibility of target disappearance.

Generally speaking, the PHD is an arbitrary function, and computation of the integrals necessary to implement the filter is a complicated if not intractable task. To simplify the problem, Vo et al. proposed approximating the PHD as a Gaussian mixture [80], and the AEGIS UKF scheme can be employed to predict and update the GMM components [11]. The GMM approximation of the posterior PHD at time t_k is given by

$$\nu_k(\mathbf{x}) \approx \sum_{j=1}^{J_k} w_k^{(j)} p_g(\mathbf{x}_k; \hat{\mathbf{x}}_k^{(j)}, P_k^{(j)}) \quad (2.55)$$

where $w_k^{(j)}$ are the weights, adhering to the property $w_k^{(j)} \geq 0$, and $\hat{\mathbf{x}}_k^{(j)}$ and $P_k^{(j)}$ denote the mean and covariance of the component. By substituting into Eq. (2.52), it is clear that in this scheme, the number of targets can be estimated as

$$N_k \approx \sum_{j=1}^{J_k} w_k^{(j)} \quad (2.56)$$

The predictor step is performed for each component using Eqs. (2.57)-(2.59), and components may be split according to the procedure outlined for AEGIS. At the end of the prediction step, the weights are multiplied by the probability of survival.

$$\bar{w}_k^{(j)} = p_S(\mathbf{x}_k^{(j)}) \bar{w}_k^{(j)} \quad (2.57)$$

$$\bar{\mathbf{x}}_k^{(j)} = \sum_{l=0}^{2L} W_l^m \bar{\chi}_{k,l}^{(j)} \quad (2.58)$$

$$\bar{P}_k^{(j)} = Q_k + \sum_{l=0}^{2L} W_l^c (\bar{\chi}_{k,l}^{(j)} - \bar{\mathbf{x}}_k^{(j)}) (\bar{\chi}_{k,l}^{(j)} - \bar{\mathbf{x}}_k^{(j)})^T \quad (2.59)$$

Computing the measurement update of the PHD function amounts to computing the updated weights, means, and covariances of the GMM components. Each component from the predictor step will produce $m_k + 1$ components from the corrector step corresponding to each of the m_k measurements and the possibility of a missed detection. This discussion considers these two possibilities separately.

In the event of a missed detection, the component means and covariances are the same as the *a priori* values. The weights are scaled by the factor $1 - p_D$, which causes the *a priori* components

to be downweighted when the probability of detection is high, but retained when the probability of detection is low.

$$w_k^{(j)} = [1 - p_D(\mathbf{x}_k^{(j)})] \bar{w}_k^{(j)} \quad (2.60)$$

$$\hat{\mathbf{x}}_k^{(j)} = \bar{\mathbf{x}}_k^{(j)} \quad (2.61)$$

$$P_k^{(j)} = \bar{P}_k^{(j)} \quad (2.62)$$

For the case where an object is assumed to be detected, a measurement update is computed for every GMM component using each measurement one at a time. The mean and covariance updates are computed similarly to the AEGIS UKF, with additional notation indicating which measurement is used. Note that throughout this dissertation, the index i is associated with measurements and j is associated with GMM components.

$$P_{zz}^{(i,j)} = R_k^{(i)} + \sum_{l=0}^{2L} W_l^c (\zeta_{k,l}^{(j)} - \bar{\mathbf{z}}_k^{(j)}) (\zeta_{k,l}^{(j)} - \bar{\mathbf{z}}_k^{(j)})^T \quad (2.63)$$

$$P_{xz}^{(j)} = \sum_{l=0}^{2L} W_l^c (\bar{\chi}_{k,l}^{(j)} - \bar{\mathbf{x}}_k^{(j)}) (\zeta_{k,l}^{(j)} - \bar{\mathbf{z}}_k^{(j)})^T \quad (2.64)$$

$$K_k^{(i,j)} = P_{xz}^{(j)} [P_{zz}^{(i,j)}]^{-1} \quad (2.65)$$

$$\hat{\mathbf{x}}_k^{(i\bar{J}_k+j)} = \bar{\mathbf{x}}_k^{(j)} + K_k^{(i,j)} (\mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^{(j)}) \quad (2.66)$$

$$P_k^{(i\bar{J}_k+j)} = \bar{P}_k^{(j)} - K_k^{(i,j)} P_{zz}^{(i,j)} [K_k^{(i,j)}]^T \quad (2.67)$$

The updated weights must be multiplied by the probability of detection. The likelihood term in Eq. (2.69) accounts for the association distance between the measurement and the *a priori* GMM component. By scaling the weight by this value, the equation will effectively downweight updates for components when the measurement is far from the predicted value, and keep the weight high when the measurement is close. The normalization enforces the standard multitarget assumption that each measurement is produced by only one target.

$$g_k^{(i,j)} = p_g(\mathbf{z}_k^{(i)}; \bar{\mathbf{z}}_k^{(j)}, P_{zz}^{(i,j)}) \quad (2.68)$$

$$w_k^{(j)} = p_D(\mathbf{x}_k^{(j)}) \frac{g_k^{(i,j)} \bar{w}_k^{(j)}}{\kappa_k + \sum_{l=1}^{\bar{J}_k} g_k^{(i,l)} \bar{w}_k^{(l)}} \quad (2.69)$$

After completing the missed detection and measurement update steps, the total number of GMM components is $\bar{J}_k \cdot (m_k + 1)$, and the full PHD function is the weighted sum computed using Eq. (2.55). However, many of the components will have small weights, and some may be very close to each other in the state space, indicating they actually represent the same object. A simple three step procedure may be incorporated to keep the number of components from growing to unmanageable numbers while retaining the accuracy and intent of the GM method [80]. First, any components with weights below a threshold T are removed. Second, if two components have a Mahalanobis distance within a merging threshold U , they are combined. Finally, if more than J_{max} components remain, only the J_{max} components with the largest weights are kept. Following this procedure, the component weights must again be normalized. The full AEGIS PHD algorithm is summarized in Appendix B.

While the process above and the result may seem like a data association algorithm, the fundamental difference is that at no point is it necessary to determine which measurement comes from which object. The filter is capable of computing updates for all GMM components using all measurements, and the resulting weights determine which components are kept. The filter does not explicitly track which measurements come from which objects; if one is close, its information will be used and the component will be kept. If the measurement is far away, the component will likely be discarded. This marks a fundamental difference from the conventional algorithms, which must know which measurement came from which object to compute the Kalman filter update. It makes the PHD filter more computationally efficient while sacrificing the knowledge of specific target IDs. Target IDs can be maintained along with the components using a separate track table, which is a relatively simple and effective ad hoc modification [57].

Extraction of object states and uncertainties is a straightforward process, though a few similar options exist. This dissertation takes each of the N_k highest peaks of the PHD function to represent an object in space, where the number of objects is estimated from Eq. (2.56). Depending on the application, the peak GMM component covariances are taken to represent the uncertainties for each object, or a cluster is formed around each peak and the weighted mixture of GMM components

in the cluster is taken to represent the object's PDF. These methods are described in more detail in the applicable sections of following chapters. Due to the truncation of information in forming the PHD and the GMM approximation, there is no guarantee that the covariance of the GMM component is an accurate reflection of the object's uncertainty. With this consideration, it is noted that in the literature [49, 80] and for the simulations included later, approximating the uncertainty using the GMM covariance works reasonably well, and for the simulations in this chapter, the PHD filter covariances generally match those reported by the MHT and JPDA filters.

2.3.2 CPHD Filter

One of the drawbacks of the PHD filter is that there can be a high variability in the expected number of targets from one epoch to the next [49]. To address this issue, the CPHD filter propagates and updates the cardinality distribution in addition to the PHD function. This produces a more computationally complex filter, but also provides a more stable estimate of the number of objects at each time.

The GM-CPHD recursion is provided below [83, 77], not including terms for target birth. The prediction step is similar to the PHD filter, and includes an extra step to predict the cardinality distribution,

$$\nu_{k|k-1}(\mathbf{x}) = p_S \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} p_g(\mathbf{x}; \bar{\mathbf{x}}_k^{(j)}, \bar{P}_k^{(j)}) \quad (2.70)$$

$$p_{k|k-1}(n) = \sum_{l=n}^{\infty} C_n^l p_{k-1}(l) p_S^n (1 - p_S)^{l-n} \quad (2.71)$$

where $C_n^l = \frac{l!}{n!(l-n)!}$ is the binomial coefficient, and the probability of survival p_S is assumed to be independent of target state for simplicity.

The update equations for the cardinality distribution and PHD are coupled, introducing

additional complexity,

$$\nu_k(\mathbf{x}) = \frac{\langle \Psi_k^1[w_{k|k-1}, Z_k], p_{k|k-1} \rangle}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{k|k-1} \rangle} (1 - p_D) \nu_{k|k-1}(\mathbf{x}) + \sum_{i=1}^{m_k} \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(\mathbf{z}^{(i)}) p_g(\mathbf{x}; \mathbf{x}_k^{(j)}, P_k^{(j)}) \quad (2.72)$$

$$p_k(n) = \frac{\Psi_k^0[w_{k|k-1}, Z_k](n) p_{k|k-1}(n)}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{k|k-1} \rangle} \quad (2.73)$$

where the angle bracket $\langle a, b \rangle$ is used to denote the inner product of a and b and the following definitions are used

$$\Psi_k^u[w, Z](n) = \sum_{j=0}^{\min(m_k, n)} (m_k - j)! p_\kappa(m_k - j) P_{j+u}^n \frac{\langle 1 - p_D, \nu \rangle^{n-(j+u)}}{\langle 1, \nu \rangle^n} \sigma_j(\Lambda_k(w, Z)) \quad (2.74)$$

$$\Lambda_k(w, Z) = \left\{ \frac{\langle 1, \kappa_k \rangle}{\kappa_k(\mathbf{z}_k^{(i)})} p_D \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} p_g(\mathbf{z}_k^{(i)}; \bar{\mathbf{z}}_k^{(j)}, P_{zz}^{(i,j)}) : \mathbf{z}_k^{(i)} \in Z \right\} \quad (2.75)$$

$$w_k^{(j)}(\mathbf{z}_k^{(i)}) = p_D w_{k|k-1}^{(j)} p_g(\mathbf{z}_k^{(i)}; \bar{\mathbf{z}}_k^{(j)}, P_{zz}^{(i,j)}) \frac{\langle \Psi_k^1[w_{k|k-1}, Z_k \setminus \{\mathbf{z}^{(i)}\}], p_{k|k-1} \rangle}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{k|k-1} \rangle} \frac{\langle 1, \kappa_k \rangle}{\kappa_k(\mathbf{z}_k^{(i)})} \quad (2.76)$$

where m_k is the number of measurements, $p_\kappa(\cdot)$ is the cardinality of clutter, $\kappa_k(\cdot)$ is the PHD of clutter, and $\sigma_j(\cdot)$ are the elementary symmetric functions, which can be computed using a formulation from Mahler [49]. The probability of detection p_D is assumed to be independent of target state. At time t_k , the number of targets can be estimated using the PHD as before, or using the estimated cardinality distribution.

$$N_k = \sum_{n=1}^{\infty} n \cdot p_k(n) \quad (2.77)$$

Within the AEGIS framework, the weights, means, and covariances of the GMM components are propagated using Eqs. (2.57) - (2.59), while the cardinality is propagated using Eq. (2.71). The update step is similar to the AEGIS PHD, where a measurement update is computed for each component using each of the m_k measurements, and the *a priori* information is retained to account for missed detections. For the case of missed detections, the weights are computed from the first term of Eq. (2.72).

$$w_k^{(j)} = \frac{\langle \Psi_k^1[w_{k|k-1}, Z_k], p_{k|k-1} \rangle}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{k|k-1} \rangle} (1 - p_D) w_{k|k-1}^{(j)} \quad (2.78)$$

For detected objects, the updated weights are computed from the second term of Eq. (2.72), which includes factors of p_D and the individual likelihood of measurement to component associations g_{ij} .

Following the measurement update, the GMM components can be pruned or merged following the same procedure as the PHD filter. The full AEGIS CPHD algorithm is included in Appendix B.

2.4 Clutter Model

Throughout this chapter, references have been made to false alarms, or clutter, which are incorporated in each filter model in slightly different ways. The MHT and JPDA filters both make use of a number density λ , which indicates the expected number of clutter returns per unit volume of the measurement space. The PHD filter requires an assumed distribution of the PHD of clutter, $\kappa(\mathbf{z})$, which is also needed for the CPHD filter. The CPHD filter further requires an assumed distribution for the cardinality of clutter.

The clutter PHD is assumed to be uniformly distributed in the sensor field of view (FOV), and can be defined as

$$\kappa(\mathbf{z}) = \lambda_{\kappa} \cdot \mathcal{U}(\mathbf{z}) \quad (2.79)$$

$$\mathcal{U}(\mathbf{z}) = \begin{cases} 1/V_s & \text{if } \mathbf{z} \in \text{FOV} \\ 0 & \text{if } \mathbf{z} \notin \text{FOV} \end{cases} \quad (2.80)$$

where V_s is the sensor volume. The cardinality of clutter is assumed to be Poisson-distributed, with the mean number of clutter returns each scan given by $\lambda_{\kappa} = \lambda \cdot V_s$.

$$p_{\kappa}(n) = \frac{(\lambda_{\kappa})^n}{n!} \exp(-\lambda_{\kappa}) \quad (2.81)$$

The Poisson assumption is standard in the multitarget literature, and it allows for derivation of the PHD and CPHD update Eqs. (2.54) and (2.72)-(2.76) [49, 80, 83]. The additional assumptions regarding the clutter PHD simplify some of the terms in these update equations.

$$\kappa_k(\mathbf{z}) = \lambda \cdot V_s \cdot 1/V_s = \lambda \quad (2.82)$$

$$\frac{\langle \mathbf{1}, \kappa_k \rangle}{\kappa_k(\mathbf{z})} = \frac{\lambda \cdot V_s \int \mathcal{U}(\mathbf{z}) d\mathbf{z}}{\lambda} = V_s \quad (2.83)$$

$$\int \mathcal{U}(\mathbf{z}) d\mathbf{z} = 1 \quad (2.84)$$

2.5 Optimal Subpattern Assignment Metric

Another concept from single target statistics that must be generalized to the multitarget problem is the miss distance, or error, between the estimated and true state. In the context of the RFS, the multitarget miss distance must provide the difference between two sets of vectors, accounting for errors both in cardinality and the state estimates of individual objects. The Optimal Subpattern Assignment (OSPA) metric addresses this need, and offers several advantages over previous multitarget error metrics [68].

OSPA is a consistent metric on the space of finite sets, and provides a meaningful physical interpretation of state and cardinality errors by assigning the largest possible subset of estimated and true states, computing the error between them, and adding a fixed error for each point that is not assigned. The computation can be summarized in three steps.

Given two random finite sets, $X = \{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_n\}$, the p -th order OSPA metric with cutoff c is defined as

$$d_p^{(c)}(X, Y) = \left[\frac{1}{n} \left(\min_{\pi \in \Pi_n} \sum_{i=1}^m d^{(c)}(\mathbf{x}_i, \mathbf{y}_{\pi(i)})^p + c^p(n - m) \right) \right]^{1/p} \quad (2.85)$$

$$d^{(c)}(\mathbf{x}_i, \mathbf{y}_{\pi(i)}) = \min(c, d(\mathbf{x}_i, \mathbf{y}_{\pi(i)})) \quad (2.86)$$

where Π_n represents the set of permutations on $\{1, \dots, n\}$ and it is assumed $m \leq n$. The term $d(\mathbf{x}_i, \mathbf{y}_{\pi(i)})$ represents a distance metric in the single target space, for instance, the Euclidean distance $\|\mathbf{x}_i - \mathbf{y}_{\pi(i)}\|_2$. For the case $m > n$, the metric can be computed as $d_p^{(c)}(Y, X)$. The steps to compute OSPA are as follows:

1. Find the optimal subpattern assignment between X and Y that minimizes the distance metric $d(\mathbf{x}, \mathbf{y})$ between m vectors. This can be done using a 2D assignment algorithm such as auction [5].
2. For each vector \mathbf{y}_i , let $\alpha_i = c$ if no assignment is made, or the minimum of c and the single target distance metric of the assignment.

3. Compute the p -th order average $d_p^{(c)}(X, Y) = ((1/n) \sum_{i=1}^n \alpha_i^p)^{1/p}$.

OSPA is useful because it allows a single number to represent a combined state and cardinality error. However, it can also be broken into separate state and cardinality components by reporting the contributions of the largest matched subset and each additional penalty term separately. The state errors can be further decomposed by considering portions of the state vector separately, which may provide additional insight into filter performance. Results in this chapter consider position and velocity errors in the state estimate, and report the estimated number of objects and GMM components used in the filter to provide a detailed description of filter performance.

2.6 Numerical Simulation

To demonstrate the performance of the filters, a test case is developed for a varying number of objects in geosynchronous orbit. The equations of motion used are those of the two-body problem.

$$\ddot{\mathbf{r}} = \frac{-\mu}{r^3} \mathbf{r} \quad (2.87)$$

where \mathbf{r} is the position vector and μ is the gravitational parameter. The state vector consists of the position and velocity in the Earth-Centered Inertial (ECI) frame.

$$\mathbf{x} = [x \quad y \quad z \quad \dot{x} \quad \dot{y} \quad \dot{z}]^T \quad (2.88)$$

To prevent the filter covariance from collapsing and causing measurements to be ignored, a simple process noise model is implemented, in which fixed values are multiplied by the time interval of propagation and added along the diagonal terms of the state covariance matrix. The σ_i values are specified in the Radial/In-track/Cross-track (RIC) directions and rotated to the ECI frame, values are provided in Table 2.6.

The measurements used are topocentric right ascension and declination taken from a single ground station located in Maui at the coordinates given in Table 2.3. The transformation between the ECI and Earth-Centered Earth-Fixed (ECEF) frames is modeled as a simple z-axis rotation using the Earth rotation angle at each observation time. The sensor is assumed to have a 2 degree

field of view in both axes, and white measurement noise is added with standard deviations provided in Table 2.5.

Table 2.3: Station Location in ECEF

Station	x_s (km)	y_s (km)	z_s (km)
Maui	-5465.210	-2403.610	2242.120

The measurement equations are given by

$$\alpha = \tan^{-1} \left(\frac{y - y_{si}}{x - x_{si}} \right) \quad \delta = \sin^{-1} \left(\frac{z - z_{si}}{\rho} \right) \quad (2.89)$$

where $\rho = \sqrt{(x - x_{si})^2 + (y - y_{si})^2 + (z - z_{si})^2}$ is the range and the subscript si denotes a ground station coordinate in ECI.

2.6.1 Test Case Description

The test case consists of estimating the states of objects in geosynchronous orbit, where the number of objects is increased from 2 - 10. In each case the first object is the EchoStar1 satellite, with starting positions for other objects randomly initialized by varying the inclination, eccentricity, and mean anomaly from the EchoStar1 elements.¹ The variations cause the objects to drift over the course of the simulation, though with the sensor tasked to point at the EchoStar1 satellite each epoch, none leave the field of view. A plot of the measurements in the field of view for the 10 object case is provided in Figure 2.4, with each color representing a different object. The satellite elements and variations are provided in Table 2.4.

A dense measurement set of observations every 10 minutes over a 12 hour window is used, including missed detections and a Poisson-distributed number of clutter returns each scan, according to the parameters in Table 2.7. For each value of the number of objects in the simulation, two instances of the test are performed using different measurements. Each filter is initialized with the correct number of GMM components, with covariances specified in Table 2.5 and initial states perturbed from the truth by these same values. The CPHD is additionally initialized with a uniform

¹ TLE obtained from www.space-track.org

Table 2.4: Initial State Parameters (EchoStar1)

Parameter	Value
a	42164.573 (km)
e	0.0002878
i	0.006 (deg)
Ω	278.657 (deg)
ω	139.8697 (deg)
M	181.4332 (deg)
σ_i	0.4 (deg)
σ_M	0.25 (deg)
σ_e	0.003

Table 2.5: Initial State and Measurement Standard Deviations

Parameter	Value
$\sigma_x = \sigma_y = \sigma_z$	10 (km)
$\sigma_{\dot{x}} = \sigma_{\dot{y}} = \sigma_{\dot{z}}$	0.01 (km/s)
$\sigma_\alpha = \sigma_\delta$	1 (arcsec)

Table 2.6: Process Noise Standard Deviations in RIC Frame

Parameter	Value
$\sigma_r = \sigma_c$	1e-11 (km)
σ_i	1e-10 (km)
$\sigma_{\dot{r}} = \sigma_{\dot{c}}$	1e-17 (km/s)
$\sigma_{\dot{i}}$	1e-16 (km/s)

Table 2.7: Filter Parameters

Parameter	Value
p_S	1
p_D	0.9
λ	2.5 (#/deg ²)
Gate	10
Murty k -best	10
N -Scan	3
MHT T	$0.01 * p_{\max}(\text{H})$
PHD T	$0.01 * w_{\max}$
PHD U	4
PHD J_{\max}	20

cardinality distribution over the range of 0 to 20 objects. Both the PHD and CPHD filters are implemented with track labels as described in Reference [57]. Additional filter parameters described

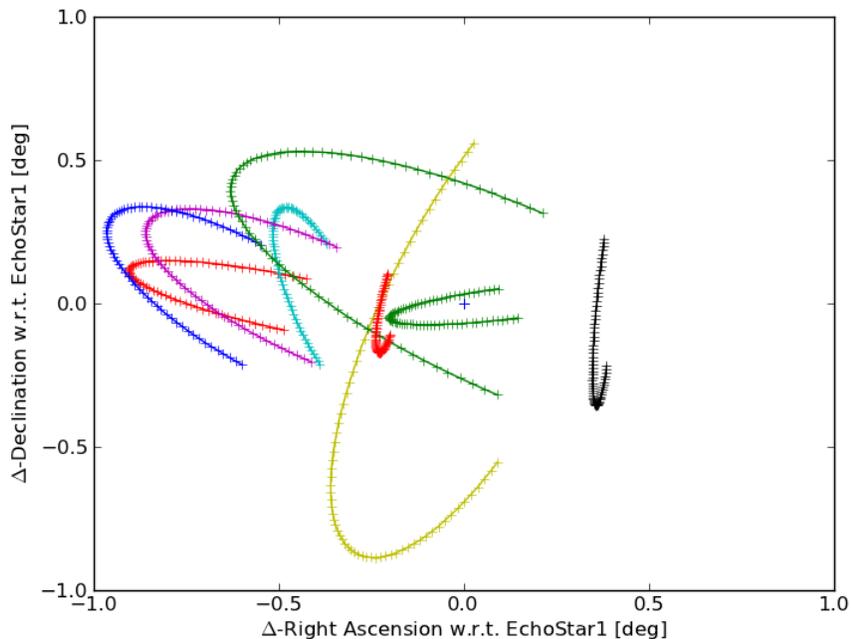


Figure 2.4: Measurement Visualization

in this chapter are provided in Table 2.7.

2.6.2 Results and Discussion

As a first analysis, the average execution times from the two measurement sets are plotted against the number of objects for all four filters, shown in Figure 2.5. The graph shows that for 2-10 objects, the PHD, CPHD, and JPDA filters all require a similar level of computational effort. The MHT filter, as implemented here, becomes markedly slower.

Beginning with 4 objects, the possibility of missed detections forces the MHT filter to generate numerous additional hypotheses, which increases the computation required. With a probability of detection less than one, missed detection hypotheses must be generated each epoch, and retained until future measurements warrant their removal per the N -scan pruning method. As a result of the gate checks discussed in Section 2.2.1, clutter is not of significant consequence, because measurements far from objects are not considered and false measurements near objects do not

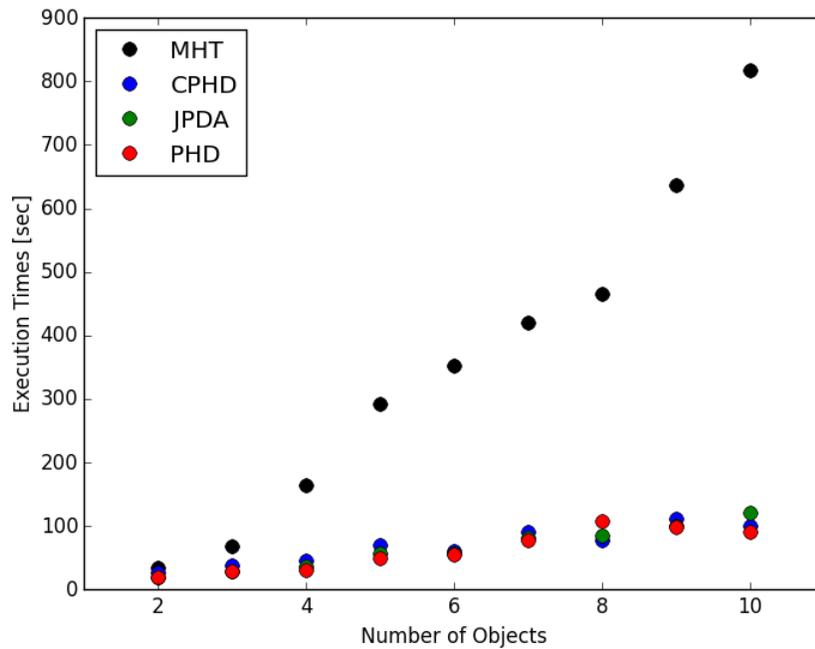


Figure 2.5: Execution Times

affect state estimates significantly.

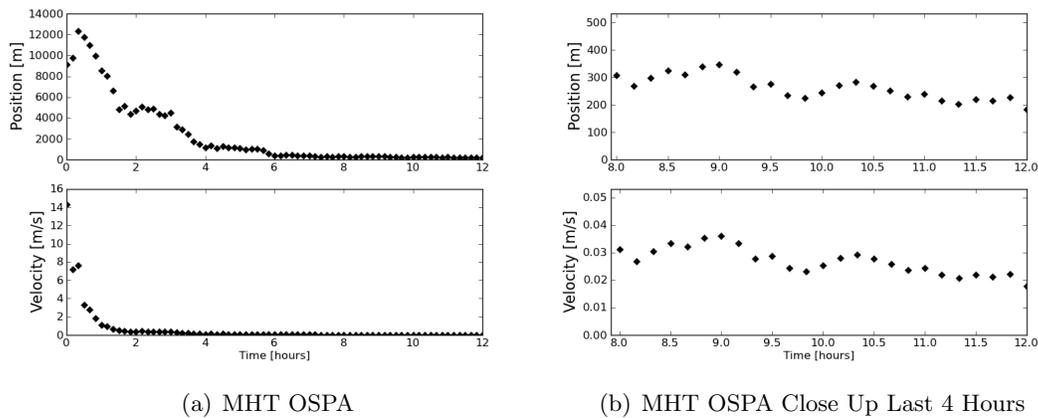


Figure 2.6: MHT 10 Object OSPA Results

To further compare filter performance, individual results are provided for one of the 10 object cases, beginning with MHT. Figure 2.6 provides the 2nd order position and velocity OSPA, computed using the Euclidean distance between the estimated target states and truth. The results

presented are for the most likely hypothesis, which in this case correctly matches all measurements to targets and is expected to provide the best achievable accuracy for any of the filters. From the figures, it is clear that the filter significantly reduces the position and velocity errors over time. The close up view shows position errors around 200 meters by the final time, which matches the level of the measurement noise projected to Cartesian coordinates at GEO.

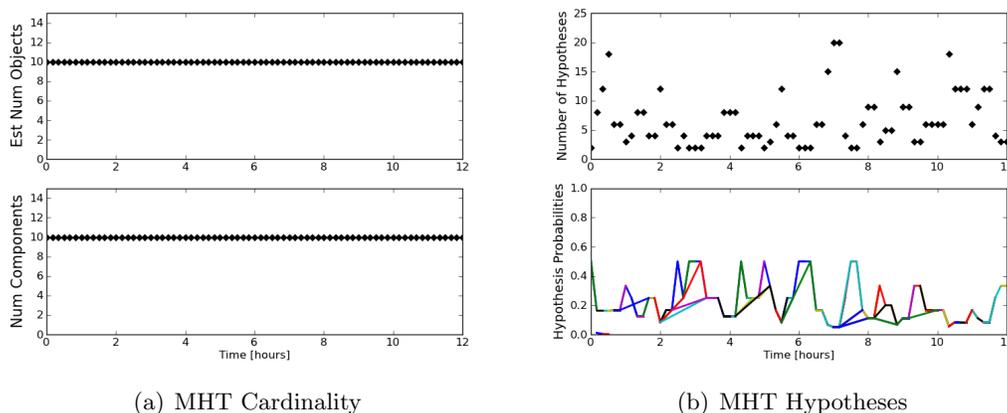


Figure 2.7: MHT 10 Object Results

Figure 2.7 provides the cardinality results and a plot of the number of hypotheses and their associated probabilities at each time. From the figure on the left, it is clear that MHT maintains the correct estimate of the number of objects, using only one component per object throughout. On the right, the plot shows the reason MHT has a large computation time for this case. The number of hypotheses fluctuates throughout the simulation, increasing to 20 at some points, which would require much more time to propagate than the single hypothesis maintained by each of the other methods. It is interesting to note that the probabilities all converge to the same value around the 10 hour mark, and only increase or decrease as the number of hypotheses changes. This indicates that the filter estimates themselves in each hypothesis are very nearly, if not entirely, identical. An additional step that could reduce these redundancies is to merge similar hypotheses, but this would add to the complexity of the filter.

This issue speaks to a general problem of MHT, namely that trying to solve a multitarget

problem with multiple single target filters requires more ad hoc modifications as the problem gets more complex [49]. Merging hypotheses is a well-established technique which would likely address the issue here and create a more efficient filter. Other techniques such as track-oriented or multi-scan MHT could also remove these redundancies and reduce computational effort. Researchers have developed a number of variations of MHT to approach problems and there is ongoing research and application of MHT filters [6, 60, 1]. In this work, some modifications are used as a reasonable point of comparison, but the objective is not to advance this research or exhaustively study available options. For the comparison included, using measurement-oriented MHT with Murty hypothesis selection and N -scan pruning, it is clear that the MHT filter will never be as efficient as techniques maintaining a single hypothesis. Missed detections will always cause new hypotheses to be created, some may never be pruned, and the filter will not converge to a single hypothesis in general without further modification.

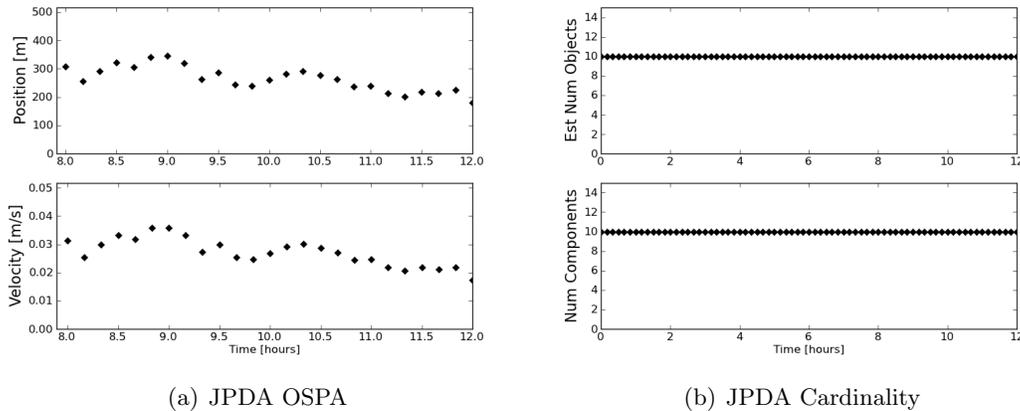


Figure 2.8: JPDA 10 Object Results

Figure 2.8 provides the results for the JPDA filter. The OSPA results shown only consider the end of the simulation window, and the filter achieves the same level of accuracy as MHT. This result confirms the above consideration that merging hypotheses produces a more efficient filter without sacrificing accuracy in the state estimate. In addition, the filter maintains the correct estimate of the number of targets throughout, using a single GMM component to represent each.

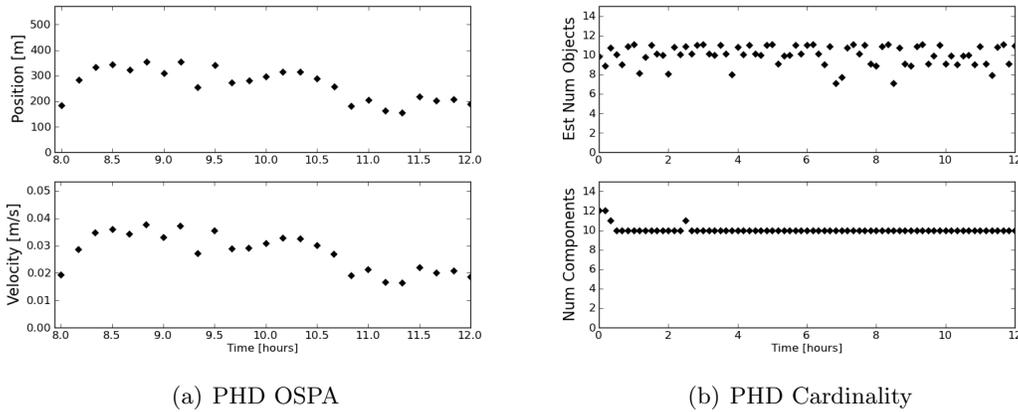


Figure 2.9: PHD 10 Object Results

Figure 2.9 provides the results for the PHD filter. The OSPA results are similar to the conventional filters and the state estimate accuracy is good. However, the estimated number of targets fluctuates significantly through the course of the simulation, varying between 7 and 12, and never settling or converging to any one value. The variability in target number is primarily driven by missed detections and is a known issue in the PHD filter. In this instance, the issue is significant enough to disqualify the PHD filter as a legitimate option for SSA problems.

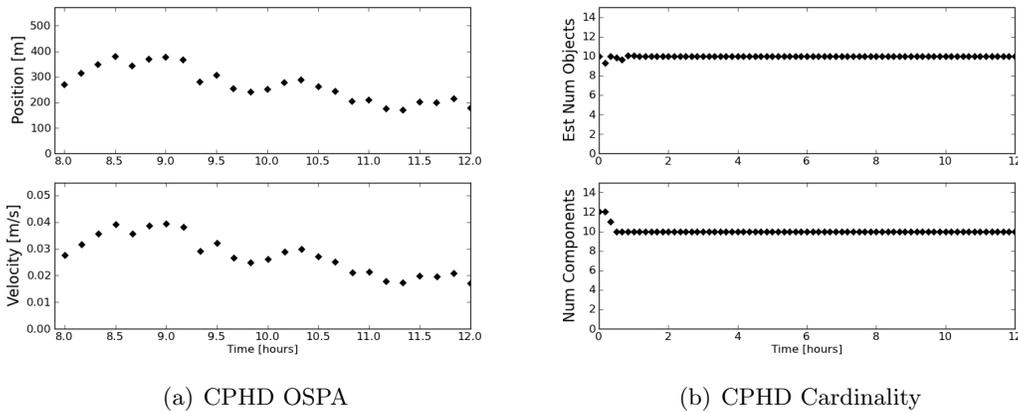


Figure 2.10: CPHD 10 Object Results

Figure 2.10 provides the results for the CPHD filter. The CPHD filter was derived specifically to address the issue of target variability in the PHD filter, and it is successful in this test. There

are some small variations in the beginning, but this is because the filter is initialized with a uniform cardinality and allowed to converge to the correct estimate of the number of targets. It should be noted that it is the only filter to begin with this uncertainty, and so the short time to converge on 10 objects should not be considered a disadvantage as compared to MHT or JPDA. In addition, the OSPA plot demonstrates that the filter performs comparably to the others in terms of the estimated state accuracy.

2.7 Chapter Summary

The goal of this study is to establish FISST as a viable approach for the SSA problem. While there are a large number of both conventional and FISST-based filters that could be implemented, this study only considers a small set of options. Issues were found with both the PHD and MHT approaches, in estimating the number of targets and computational burden respectively. The CPHD and JPDA filters may be considered variations to these filters, addressing the issues as discussed. Both filters achieved similar performance in term of state estimation accuracy and computational burden.

In a more qualitative sense, there are many similarities between conventional and FISST-derived multitarget filters. By virtue of combining information from the full measurement set into a single update for all objects, the PHD filter is philosophically similar to JPDA. In fact, recent work has established that a variant of PDA, known as Integrated Probabilistic Data Association (IPDA), can be derived from FISST, and doing so provides a more rigorous process to model target birth and death [43, 9]. Other researchers have developed variants of multi-Bernoulli filters using FISST [82, 81], which closely resemble track-oriented MHT. This overlap is somewhat expected as the filters are all Bayesian estimators, and underlying assumptions of Gaussian distributions and the use of the unscented transform to propagate and rectify uncertainties are commonly applied. The mathematical formalism of FISST does offer an advantage in making principled approximations to derive filters with a specific purpose, such as the CPHD to address the poor estimate of target number produced by the PHD. In reviewing the literature, no cardinalized equivalent of JPDA was

found to match the CPHD. Ultimately, the CPHD filter appears to be a suitable approach for GEO SSA. The remainder of this dissertation considers application of the CPHD filter to more advanced problems in SSA, beginning with the allocation of sensor resources.

Chapter 3

Information Theoretic Sensor Allocation

One of the primary challenges in space situational awareness is tracking a large number of objects using a limited number of sensors. The sensor limitation leads to sparse data scenarios, in which long gaps occur between measurements of individual objects. Selection of an appropriate sensor allocation (SA) scheme is essential to maintaining an accurate catalog of resident space objects (RSOs). A good sensor allocation scheme will yield accurate state estimates and a consistent representation of the uncertainty associated with each object, while additionally preventing the loss of known objects from the catalog.

Sensor allocation schemes can be organized in a top-down, information theoretic manner, or as a bottom-up composition of heuristic rules [50]. The first method is mathematically rigorous and can be used to efficiently collect measurements on the full multitarget system, but it is also more computationally intensive. The latter is simpler but may not achieve the desired performance in estimation accuracy and catalog maintenance.

Schemes can further be classified in terms of temporal assignment as single-step or multistep. Single-step algorithms use predicted information for the next time step to make a selection of which targets to observe, while multistep systems allocate sensors through a future time window and therefore account for targets entering and leaving the sensor field of regard (FOR). Multistep approaches are generally more effective, provided target dynamics are well-known and targets are not maneuvering unexpectedly [50].

A final classification of the sensor allocation approach is based on the objective function, for

which many possibilities exist. One of the simplest is to maximize the posterior expected number of targets (PENT), or the related PENTL, which accounts for targets of interest [46, 44]. PENT was originally developed for use with the PHD filter [45, 80]. Higher order approximate filters such as the CPHD [48, 83] and Cardinality Balanced Multitarget Multi-Bernoulli (CB-MeMBer) [84] maintain an estimate of the cardinality distribution, and therefore admit more advanced SA objective functions, such as regional variance [15] or cardinality variance [33]. While PENT seeks to maximize the estimated number of targets, minimizing the estimated cardinality variance translates to an expected minimization in the estimated cardinality error. The philosophy of both approaches is similar. In the case that measurements provide sufficient information on object states, driving to an accurate estimate of target number produces a correspondingly accurate estimate of individual target states [33]. However, in the case of angles-only observations of GEO space objects, the measurements do not provide direct information on the velocity or radial position states of the object. Furthermore, neither PENT nor cardinality variance directly account for state uncertainties, opening up the potential for better objective functions for GEO SSA based on information theory.

Information theory provides a means to quantify the difference between the prior and posterior multitarget PDFs, such that maximizing the information gain functional produces the greatest reduction in posterior uncertainty. A number of information functionals fall in the category of Csiszár discrimination [89, 50], including the whole class of Rényi α -divergence [64], Kullback-Leibler (KL) cross-entropy [40], Hellinger discrimination [89], and Shannon entropy [70]. In fact, the Rényi divergence reduces to KL in the limit as $\alpha \rightarrow 1$ [30], or Hellinger discrimination for the choice $\alpha = 0.5$ [66, 30]. With additional and significant approximation, PENT can be shown to derive from KL entropy [50]. An additional information functional outside the category of Csiszár discrimination is the Cauchy-Schwarz divergence [32, 16]. This chapter considers the use of both the Rényi and Cauchy-Schwarz divergences as potential information gain functionals to perform sensor tasking, ultimately selecting the Rényi divergence as formulated for the CPHD filter for application to the GEO SSA problem.

The remainder of this chapter is organized as follows. Section 3.1.1 summarizes the previously

discussed CPHD filter, extending the probability of detection calculation for cases where objects leave the sensor field of view. Section 3.1.2 provides the details of the information gain functionals and examines the use of both the Rényi and Cauchy-Schwarz divergences as the reward function in a simplified sensor tasking scenario. Section 3.1.3 provides an overview of the sensor tasking methodology. A detailed SSA simulation is included in Section 3.2, in which a large number of geosynchronous objects are scheduled to be tracked by a small number of ground-based sensors.

3.1 Background

3.1.1 Review of CPHD Filter

The GM-CPHD recursion equations are provided for reference below. In this chapter, target birth is not considered and is once again omitted from the prediction equations. Following the development of Vo et al. [83], the PHD at time t_k is approximated as a weighted sum of J_k Gaussian PDFs,

$$\nu_k(\mathbf{x}) \approx \sum_{j=1}^{J_k} w_k^{(j)} p_g(\mathbf{x}; \mathbf{x}_k^{(j)}, P_k^{(j)}) \quad (3.1)$$

where $w_k^{(j)} > 0$ are the weights and $p_g(\mathbf{x}; \mathbf{x}_k^{(j)}, P_k^{(j)})$ is a multivariate Gaussian PDF with mean $\mathbf{x}_k^{(j)}$ and covariance $P_k^{(j)}$. The prediction step for the PHD and cardinality is given by

$$\nu_{k|k-1}(\mathbf{x}) = p_S \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} p_g(\mathbf{x}; \mathbf{x}_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}) \quad (3.2)$$

$$p_{k|k-1}(n) = \sum_{l=n}^{\infty} C_n^l p_{k-1}(l) p_S^n (1 - p_S)^{l-n} \quad (3.3)$$

where p_S is the probability of target survival, assumed to be independent of the target state, and $C_j^l = \frac{l!}{j!(l-j)!}$ is the binomial coefficient.

The update equations for the cardinality distribution and PHD are coupled,

$$\nu_k(\mathbf{x}) = \frac{\langle \Psi_k^1[w_{k|k-1}, Z_k], p_{k|k-1} \rangle}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{k|k-1} \rangle} (1 - p_D) \nu_{k|k-1}(\mathbf{x}) + \sum_{\mathbf{z} \in Z_k} \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(\mathbf{z}) p_g(\mathbf{x}; \mathbf{x}_k^{(j)}, P_k^{(j)}) \quad (3.4)$$

$$p_k(n) = \frac{\Psi_k^0[w_{k|k-1}, Z_k](n) p_{k|k-1}(n)}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{k|k-1} \rangle} \quad (3.5)$$

where Z_k is the measurement set, the angle bracket $\langle a, b \rangle$ is used to denote the inner product of a and b , and additional terms are given as follows:

$$\Psi_k^u[w, Z](n) = \sum_{j=0}^{\min(m_k, n)} (m_k - j)! p_\kappa(m_k - j) P_{j+u}^n \frac{\langle 1 - p_D, \nu \rangle^{n-(j+u)}}{\langle 1, \nu \rangle^n} \sigma_j(\Lambda_k(w, Z)) \quad (3.6)$$

$$\Lambda_k(w, Z) = \left\{ \frac{\langle 1, \kappa_k \rangle}{\kappa_k(\mathbf{z})} p_D \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} p_g(\mathbf{z}; \mathbf{z}_k^{(j)}, P_{zz}^{(j)}) : \mathbf{z} \in Z \right\} \quad (3.7)$$

$$w_k^{(j)}(\mathbf{z}) = p_D w_{k|k-1}^{(j)} p_g(\mathbf{z}; \mathbf{z}_k^{(j)}, P_{zz}^{(j)}) \frac{\langle \Psi_k^1[w_{k|k-1}, Z_k \setminus \{\mathbf{z}\}], p_{k|k-1} \rangle \langle 1, \kappa_k \rangle}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{k|k-1} \rangle \kappa_k(\mathbf{z})} \quad (3.8)$$

where m_k is the number of measurements, $p_\kappa(\cdot)$ is the cardinality of clutter, $\kappa_k(\cdot)$ is the PHD of clutter, $\sigma_j(\cdot)$ are the elementary symmetric functions, and $\mathbf{z}_k^{(j)}$ and $P_{zz}^{(j)}$ are the predicted mean and covariance of component j in the measurement space. The probability of detection, p_D , is assumed to be independent of the target state; however, this assumption will be revisited in the following section. At time t_k , the number of targets can be estimated using the PHD or the estimated cardinality distribution.

$$N_k = \int \nu_k(\mathbf{x}) d\mathbf{x} \approx \sum_{j=1}^{J_k} w_k^{(j)} \quad (3.9)$$

$$N_k = \sum_{n=1}^{\infty} n \cdot p_k(n) \quad (3.10)$$

The PHD update Eq. (3.4) contains two terms, the first of which accounts for the possibility of missed detections. In this case, the filter reweights the *a priori* PHD function. In the GMM approximation, the *a priori* component means and covariances are kept and are not updated by any measurement, and the updated component weights are computed from

$$w_k^{(j)} = \frac{\langle \Psi_k^1[w_{k|k-1}, Z_k], p_{k|k-1} \rangle}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{k|k-1} \rangle} (1 - p_D) w_{k|k-1}^{(j)} \quad (3.11)$$

This term produces $J_{k|k-1}$ GMM components whose weights are scaled by a factor including $(1 - p_D)$. The second term of Eq. (3.4) accounts for detected objects, and computes a measurement update for each GMM component using each measurement, thereby producing $J_{k|k-1} \cdot m_k$ components. The updated weights for these components are scaled by a factor including p_D and the individual

likelihood of measurement to component association $p_g(\mathbf{z}; \mathbf{z}_k^{(j)}, P_{zz}^{(j)})$. The net result of the measurement update step is the creation of $J_{k|k-1} \cdot (m_k + 1)$ components, many of which will have small weights and not contribute significantly to the GMM approximated PHD. To keep the problem computationally tractable, components are removed or merged based on user-defined thresholds [80].

The individual component means, covariances, and predicted measurements in Eqs. (3.2)-(3.8) are computed using the unscented transform [36, 85], as described in Chapter 2. In addition to the GMM components of the PHD, the filter maintains a list of object identifiers from the public Two-Line Element (TLE) catalog. Anytime a new component is added to the GMM as a result of the measurement update, the identifier of the original component is added to the list as well. The approach does not ensure a rigorous treatment of unique track labels incorporated in the multitarget state [50], but in the case that no target birth or death is considered, it does provide a practical means to reconstruct the time history of components and compute the average state errors for each object as described in Section 3.2.3.

In this chapter, the cardinality is initialized using a binomial distribution [49],

$$p_0(n) = C_n^J (1 - q)^J \left(\frac{q}{1 - q} \right)^n \quad (3.12)$$

where J is the number of terms in the cardinality distribution and q is the individual probability of existence for each object. If the true number of objects is N and the distribution is initialized with the value $q = N/J$, then $\underset{n}{\operatorname{argmax}} p_0(n) = N$ and the estimated number of targets computed from Eq. (3.10) is $N_k = N$.

3.1.1.1 Probability of Detection Calculation

Computing the correct probability of detection plays an important role in maintaining custody of objects in the filter. The previous section applies a constant p_D for all components, but this is clearly problematic for objects that are outside the sensor field of view (FOV) at a particular measurement time. In the case that objects are known to be outside the FOV and cannot be

detected, assigning the same p_D as for objects in the FOV will quickly cause GMM components to be downweighted and knowledge of the object to be lost. Meaningful values of p_D need to be assigned for each object, properly accounting for their potential positions with respect to the FOV.

A simple approach to determining $p_D^{(j)}$ for an individual component is to use the predicted measurement and an indicator function [42]. The probability of detection can be modeled as a product of two terms, a constant $p_{D,\text{sensor}}$ that accounts for the sensor's imperfect ability to detect objects in the FOV, and a state-dependent term

$$p_{D,\text{FOV}}^{(j)} = \begin{cases} 1 & \text{if } \mathbf{z}^{(j)} \in \text{FOV} \\ 0 & \text{if } \mathbf{z}^{(j)} \notin \text{FOV} \end{cases} \quad (3.13)$$

where $\mathbf{z}^{(j)}$ is the predicted measurement for component j computed using the unscented transform. Assuming the two processes are independent, the overall $p_D^{(j)} = p_{D,\text{sensor}} \cdot p_{D,\text{FOV}}^{(j)}$.

This approach addresses the issue, but may not be sufficient in cases where the object is near the edge of the field of view. Consider the case illustrated in Fig. 3.1(a), in which a predicted measurement lies in the FOV, but the actual object is not visible. In the limiting case that $p_{D,\text{sensor}} = 1$, the missed detection component weight will be zero and the filter will lose custody of the object. Even in the case $p_{D,\text{sensor}} < 1$, if the object remains at the edge of the FOV for several measurement epochs, the component will be downweighted significantly and could be lost.

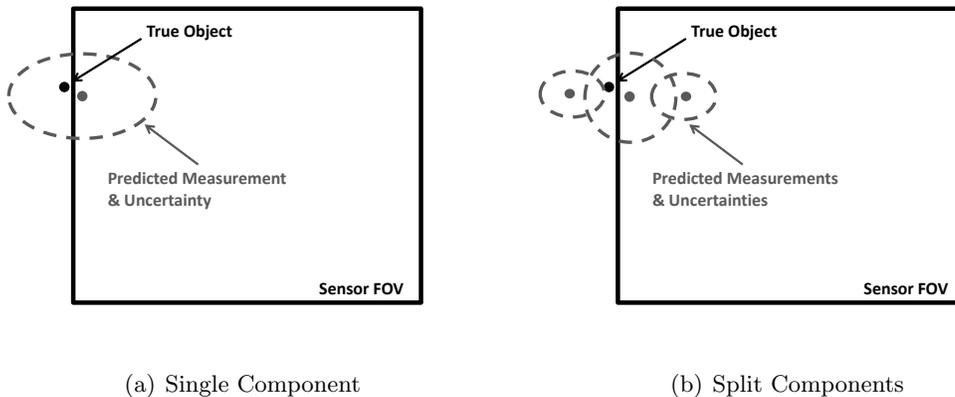


Figure 3.1: Illustration of Sensor Field of View

A solution to this issue is to split GMM components near the edge of the FOV into several new components prior to computing the probability of detection, as illustrated in Fig. 3.1(b). This allows the means of some components to lie inside the FOV and others outside, thereby accounting for the possibility of a missed detection while still maintaining a high p_D in the FOV. A simple test for splitting is to integrate the component PDF and check against a threshold, for example:

$$0.05 \leq \int_{\text{FOV}} p_g(\mathbf{z}; \mathbf{z}_k^{(j)}, P_{zz}^{(j)}) d\mathbf{z} \leq 0.95 \quad (3.14)$$

where $\mathbf{z}_k^{(j)}$ and $P_{zz}^{(j)}$ are the predicted mean and covariance in the measurement space computed using the unscented transform. If the integral is between the inequalities, the component is taken to be partially in the FOV and split according to a pre-determined library such as that used for AEGIS [18]. Having determined a value of p_D for each component, minor changes must be applied in the measurement update equations above. In Eqs. (3.4) and (3.8), the appropriate $p_D^{(j)}$ must be used for each component, more easily seen in Eq. (3.11) for the missed detection case. In Eq. (3.7), the $p_D^{(j)}$ term must be moved inside the summation, and finally the modification to Eq (3.6) is most easily simplified by writing out the inner product for the GMM approximation.

$$\langle 1 - p_D^{(j)}, \nu \rangle = \sum_{j=1}^J (1 - p_D^{(j)}) w^{(j)} \quad (3.15)$$

3.1.2 Development of Information Gain Equations

3.1.2.1 Methods

This section examines the use of Rényi α -divergence and Cauchy-Schwarz divergence as they apply to the SSA sensor allocation problem. The general form of the Rényi divergence is given by [66],

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \log \int f_1(X; \mathbf{u})^\alpha f_0(X)^{1-\alpha} dX \quad (3.16)$$

where \mathbf{u} is the sensor control vector, and $f_0(\cdot)$ and $f_1(\cdot)$ denote the prior and posterior multitarget PDFs, respectively. For the CPHD filter, the GMM approximation of the intensity function yields

the following solution, see Appendix C for details.

$$\nu_0(\mathbf{x}) \approx \sum_{i=1}^{J_0} w_i p_g(\mathbf{x}; \mathbf{m}_i, P_i) \quad \nu_1(\mathbf{x}; \mathbf{u}) \approx \sum_{j=1}^{J_1} w_j p_g(\mathbf{x}; \mathbf{m}_j, P_j) \quad (3.17)$$

$$R_C(\mathbf{u}) \approx -2 \log \sum_{n \geq 0} \left(\frac{p_1(n; \mathbf{u})}{N_1^n} \right)^{1/2} \left(\frac{p_0(n)}{N_0^n} \right)^{1/2} \cdot \left[\int \left(\sum_{i=1}^{J_0} \sum_{j=1}^{J_1} w_i w_j K_{i,j} p_g(\mathbf{x}; \mathbf{m}_{i,j}, P_{i,j}) \right)^{1/2} d\mathbf{x} \right]^n \quad (3.18)$$

$$\begin{aligned} K_{i,j} &= p_g(\mathbf{m}_i; \mathbf{m}_j, P_i + P_j) \\ P_{i,j} &= [P_i^{-1} + P_j^{-1}]^{-1} \\ \mathbf{m}_{i,j} &= P_{i,j} [P_i^{-1} \mathbf{m}_i + P_j^{-1} \mathbf{m}_j] \end{aligned}$$

where the subscripts 0 and 1 refer to the prior and posterior distributions respectively, and the value $\alpha = 0.5$ has been used as it provides the best discrimination between PDFs [66, 30]. In the case of the PHD filter, the multitarget PDF is modeled as a Poisson process, and the corresponding Rényi divergence can be derived from Eq. (3.18) by substituting a Poisson cardinality distribution $p(n) = \frac{e^{-\lambda} \lambda^n}{n!}$ and noting that in the PHD filter the Poisson mean number of targets $\lambda = N$.

$$R_P(\mathbf{u}) \approx \sum_{i=1}^{J_0} w_i + \sum_{j=1}^{J_1} w_j - 2 \int \left(\sum_{i=1}^{J_0} \sum_{j=1}^{J_1} w_i w_j K_{i,j} p_g(\mathbf{x}; \mathbf{m}_{i,j}, P_{i,j}) \right)^{1/2} d\mathbf{x} \quad (3.19)$$

While the final equation is simpler, note that both the PHD and CPHD Gaussian mixture forms of Rényi divergence require numerical integration.

As an alternative, the Cauchy-Schwarz divergence admits a closed-form solution in the Gaussian mixture PHD approximation. The general form of the Cauchy-Schwarz divergence can be written as [16]

$$C(\mathbf{u}) = \frac{1}{2} \log \int f_0(X) f_0(X) dX + \frac{1}{2} \log \int f_1(X; \mathbf{u}) f_1(X; \mathbf{u}) dX - \log \int f_0(X) f_1(X; \mathbf{u}) dX \quad (3.20)$$

Applying the GMM approximation gives the closed form solution for the PHD [16, 32].

$$C_P(\mathbf{u}) \approx \frac{1}{2} \sum_{i=1}^{J_0} \sum_{j=1}^{J_0} w_{0i} w_{0j} K_{0i,0j} + \frac{1}{2} \sum_{i=1}^{J_1} \sum_{j=1}^{J_1} w_{1i} w_{1j} K_{1i,1j} - \sum_{i=1}^{J_0} \sum_{j=1}^{J_1} w_{0i} w_{1j} K_{0i,1j} \quad (3.21)$$

No closed-form solution exists in the literature for the CPHD Cauchy-Schwarz divergence, therefore subsequent discussion and comparison of the methods is confined to the PHD equations.

3.1.2.2 Hedging

The previous section makes use of a simplified notation that requires some further discussion. The information functional $R(\mathbf{u})$ or $C(\mathbf{u})$ is also a function of the prior and posterior multitarget PDFs in the measurement space. When computing the functional, the prior PDF $f_0(X)$ is propagated to the desired time t_k . A candidate posterior PDF is generated from the CPHD update equations using a measurement set Z_k , which is dependent on the sensor control vector \mathbf{u} . The information functional, in the case of Rényi divergence, is therefore denoted as $R(\mathbf{u}, f_0(X), Z_k(\mathbf{u}))$, with a similar form for Cauchy-Schwarz. The desire is to select a sensor control vector that maximizes the information gain; however, the realized measurement set $Z_k(\mathbf{u})$ is unknown until this selection is made. The unknown measurement set can be eliminated from the information functional by taking the expected value with respect to the measurement set $Z_k(\mathbf{u})$, a process known as hedging [50].

$$R(\mathbf{u}, f_0(X)) = \mathbb{E}[R(\mathbf{u}, f_0(X), Z_k(\mathbf{u}))] \quad (3.22)$$

The expected value can be computed using multiple representative samples of measurement sets $Z_k(\mathbf{u})$ based on the given clutter intensity, probability of detection, and measurement likelihood. A simpler approach, used in this work and elsewhere in the literature [66, 3, 33], is to compute Eq. (3.22) using a single sample in which $Z_k(\mathbf{u})$ is taken to be the predicted ideal measurement set (PIMS) [44]. In this case, a single measurement set is computed assuming no clutter or measurement noise and applying $p_D = 1$ for all objects in the FOV, as determined from the component means. The PIMS measurement set is therefore

$$Z_k^{\text{PIMS}}(\mathbf{u}) = \bigcup_{\mathbf{z}_k^{(j)} \in \text{FOV}} \{\mathbf{z}_k^{(j)}\} \quad (3.23)$$

Use of the PIMS measurement set is an approximation to the expected value [50], such that

the expected information gain is given by

$$R(\mathbf{u}, f_0(X)) = \mathbb{E}[R(\mathbf{u}, f_0(X), Z_k(\mathbf{u}))] \cong R(\mathbf{u}, f_0(X), Z_k^{\text{PIMS}}) \quad (3.24)$$

The approximation is applied by simulating a measurement at the mean of each of the N_{FOV} highest weighted components. For example, if the FOV contains 10 components representing 3 estimated objects, then 3 measurements are simulated and located at the 3 largest weighted component means. The use of the highest weighted components is consistent with the approach used for state extraction in Section 3.2.3 and in the literature [83]. The updated cardinality and PHD are computed using Eqs. (3.4)-(3.5), and the Rényi divergence is computed numerically using Monte Carlo integration.

3.1.2.3 Analysis of Methods

Generally, the Cauchy-Schwarz divergence provides an attractive option for sensor allocation because it yields an analytic solution in the GMM approximation. Previous research has examined the use of the functional in instances where all objects are visible and the reward function is used to determine the ideal location of the sensor to best detect and track all objects [32, 3]. However, this is fundamentally different from the angles-only GEO SSA problem, in which only a select few objects are visible in the FOV for a given sensor pointing task, and the purpose of the sensor allocation scheme is to determine the best object or set of objects to observe at each time.

In order to better understand the behavior of the proposed information gain functionals, the single target problem is considered, for which an analytic solution of the Rényi divergence is available. The full derivation of these forms is given in Appendix C. Assuming an object represented by a single Gaussian component such that the prior and posterior weights are unity, $w_0 = w_1 = 1$, and noting that under the PIMS approximation no adjustment is made to the mean, $\mathbf{m}_0 = \mathbf{m}_1$,

the divergence functions above reduce to

$$R_C \approx -2 \log \sum_{n \geq 0} p_1(n)^{1/2} p_0(n)^{1/2} \left(\frac{|4P_{0,1}|}{|P_0 + P_1|} \right)^{n/4} \quad (3.25)$$

$$R_P \approx 2 - 2 \left(\frac{|4P_{0,1}|}{|P_0 + P_1|} \right)^{1/4} \quad (3.26)$$

$$C_P \approx \frac{1}{2} \left(\frac{1}{|4\pi P_0|} \right)^{1/2} + \frac{1}{2} \left(\frac{1}{|4\pi P_1|} \right)^{1/2} - \left(\frac{1}{|2\pi(P_0 + P_1)|} \right)^{1/2} \quad (3.27)$$

where the subscripts 0 and 1 denote the prior and posterior distributions respectively, and $P_{0,1}$ is the covariance of the product of the PDFs computed from Eq. (3.18). Note that for each formula, the minimum value 0 is found when $P_1 = P_0$, with the additional requirement of $p_1(n) = p_0(n)$ for the CPHD Rényi divergence.

To examine the behavior and relative performance of the proposed information gain functions, a simple tasking scenario is considered in which the sensor must select one out of two objects to observe at each time, see Fig. 3.2. The objects are separated such that the initial Mahalanobis distance is 10, therefore only one object is considered to be in the field of view at each time and the single target information gain equations are applied. At each time, the sensor selects one of the two objects to observe and computes a bearing measurement and corresponding covariance update using the Conventional Kalman Filter update [74]. The selection is based on maximizing the expected information gain for the single object observed, as computed using the PHD form of either the Rényi or Cauchy-Schwarz divergence, Eqs. (3.26)-(3.27).

The objects are modeled as moving with constant velocity, with state vector given by $\mathbf{x} = [x \ y \ \dot{x} \ \dot{y}]^T$. Over the course of 20 time steps, the objects move past a sensor, which computes bearing measurements $\alpha = \tan^{-1} \left(\frac{y - y_s}{x - x_s} \right)$. The covariance is propagated using a state transition matrix, and the hypothetical measurement update is computed assuming a measurement noise covariance $R_k = 1 \text{ deg}^2$.

Both objects are initialized with the same starting covariance and at all times are the same distance from the sensor, with corresponding equal magnitude measurement α . The desired sensor management behavior is to alternately select each object, thereby producing an equal number of

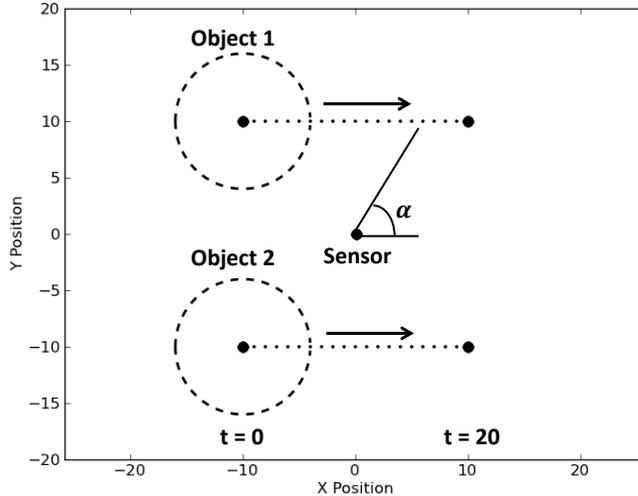


Figure 3.2: Information Gain Test Case

measurements for each and driving the final uncertainties to similar values.

Figure 3.3 provides the trace of the covariance for each object, as well as the object ID observed at each time step. The Rényi based tasking scheme performs as desired, alternately observing each object and reducing the covariances. However, the Cauchy-Schwarz scheme only gathers observations on object 1, and the covariance of object 2 grows significantly. The reason becomes clear from the form of Eq. (3.27), in which each term contains the determinant of the covariance in the denominator. When comparing the expected information gain for two objects with different initial covariances, provided the relative change as a result of the measurement update is the same for both objects, the one with the smaller initial covariance will also have the smaller final covariance and yield a larger value for the Cauchy-Schwarz divergence. Conversely, the Rényi divergence in Eq. (3.26) contains covariance terms in the numerator and denominator, which ensures the objective function will favor measurements that produce the greatest relative change in the covariance regardless of the magnitude of the initial covariance. The performance of the tasking schemes reflects that noted in previous research comparing sensor tasking based on the Fisher information matrix, a measure of absolute information gain, and Shannon entropy, a measure of relative information gain [87, 86]. In fact, the Cauchy-Schwarz divergence is similar in

form to the equation for Fisher information gain [87], and Rényi divergence and Shannon entropy are both used to quantify relative information gain, therefore these findings are consistent with previous results.

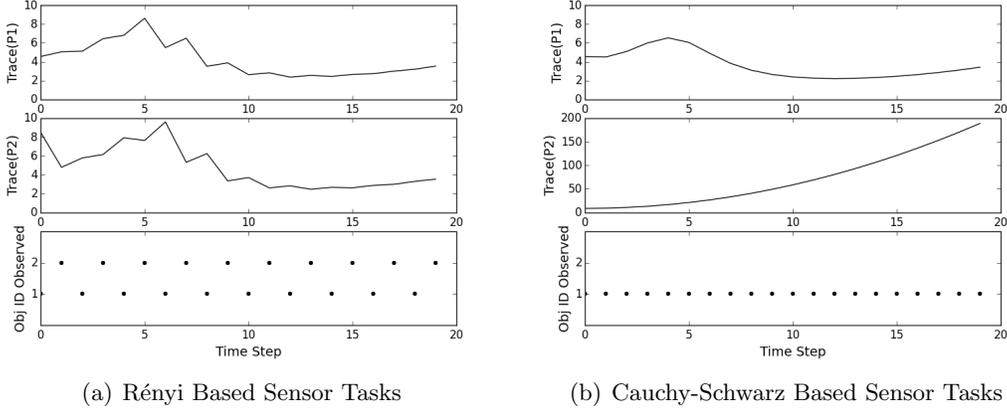


Figure 3.3: Information Gain Test Results

It should be noted that the Cauchy-Schwarz performance is highly dependent on the conditions of the test. In this simulation, a velocity of 1 m/s is used with an uncertainty of $\sigma = 0.5$ m/s and 1 second intervals between observations. Reducing the velocity, its associated uncertainty, or the time interval between observations can all result in a simulation in which the Cauchy-Schwarz based tasking alternately observes each object. The Rényi divergence shows no such dependencies, and its performance is not affected or degraded by any similar changes. It is clear that Cauchy-Schwarz is a suitable reward function for cases where all objects are visible, as the expected information gain for each object is compared against itself under different sensor conditions. However, when attempting to compare the information gain of two different objects with different initial covariances, there is potential for the Cauchy-Schwarz divergence to favor the object with the smaller starting covariance, which is generally antithetical to desired sensor tasking. Initial testing in an angles-only GEO SSA simulation demonstrated similar behavior, therefore the Rényi divergence is selected as the tasking reward function for the simulation in Section 3.2.

3.1.3 Sensor Tasking Definition

3.1.3.1 Coordinate Frames and Sensor Parameters

When considering methods to define potential sensor control vectors \mathbf{u} , several coordinate systems are of practical importance. Sensor constraints are often specified in azimuth and elevation angles, which are defined in the local East-North-Up (ENU) frame. The elevation mask ensures objects are sufficiently above the horizon to avoid obstructions and atmospheric attenuation, and certain sensors may also be restricted in the range of azimuth angles they can reach. Since many GEO objects are located in near-equatorial orbits, it can also be advantageous to provide a constraint on values of geodetic latitude, to restrict potential sensor pointing tasks to regions with a high number of targets.

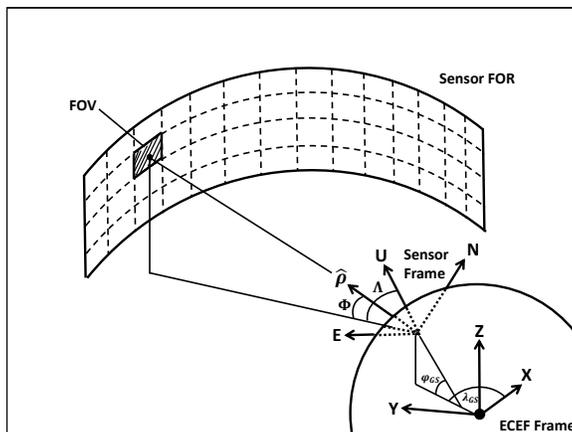


Figure 3.4: Sensor Tasking Grid

The simplest way to define potential pointing assignments is to use known object locations, provided they are visible and do not violate any sensor constraints. Another method is to divide the sensor field of regard into a grid of evenly spaced bins, such that the center of each represents a potential assignment and the field of view determines the area covered, as shown in Fig. 3.4. Defining the grid with no gaps or overlaps produces a minimal set of potential tasks to cover the entire FOR. A grid defined in geodetic latitude and longitude will be distorted in ENU coordinates unless the sensor is located at the equator, therefore a local spherical coordinate system is defined in the

ENU frame. The angle Λ is measured east from the N-U plane, and the angle Φ is measured north from the E-U plane. An even spacing of these angles produces the desired grid, and coordinates can be mapped using Eqs. (3.28)-(3.31) to check against constraints.

The pointing task is defined by the line of sight (LOS) unit vector, $\mathbf{u} = \hat{\boldsymbol{\rho}}$, which can be defined in the ENU frame can using either (Λ, Φ) or (az, el) angles.

$$\hat{\boldsymbol{\rho}}_{ENU} = \begin{bmatrix} \cos(\Phi) \sin(\Lambda) \\ \sin(\Phi) \\ \cos(\Phi) \cos(\Lambda) \end{bmatrix} = \begin{bmatrix} \cos(el) \sin(az) \\ \cos(el) \cos(az) \\ \sin(el) \end{bmatrix} \quad (3.28)$$

To compute the geodetic latitude and longitude, the position vector in Earth-Centered Earth-Fixed (ECEF) coordinates is found by solving Eq. (3.29) for a value ρ that produces a geosynchronous orbit radius $r = 42164$ km. The angles are then determined from Eqs. (3.30)-(3.31) [79],

$$\mathbf{r}_{ECEF} = \mathbf{q}_{ECEF} + \rho \hat{\boldsymbol{\rho}}_{ECEF} \quad (3.29)$$

$$\tan \lambda = \frac{y}{x} \quad \tan \phi = \frac{z}{p_e} \left(1 - e_e^2 \frac{N_e}{N_e + h} \right) \quad (3.30)$$

$$p_e = \sqrt{x^2 + y^2} \quad N_e = \frac{a_e}{(1 - e_e^2 \sin^2 \phi)^{1/2}} \quad h = \frac{p_e}{\cos \phi} - N_e \quad (3.31)$$

where \mathbf{q}_{ECEF} is the ground station position, and x , y , and z are the components of \mathbf{r}_{ECEF} .

Figure 3.5 provides a sample FOR grid for the Maui Ground-based Electro-Optical Deep Space Surveillance (GEODSS) sensor. The local (Λ, Φ) coordinates are staggered to produce a range of geodetic latitude values between ± 10 degrees. This reduced FOR grid is used for the box scan tasking mode described in the following section, and achieves the desired sensor constraints using a little over 1500 bins.

Three ground stations are modeled using the parameters provided in Table 3.1 [79]. The measurements used are topocentric right ascension and declination,

$$\alpha = \tan^{-1} \left(\frac{y - y_{si}}{x - x_{si}} \right) \quad \delta = \sin^{-1} \left(\frac{z - z_{si}}{\rho} \right) \quad (3.32)$$

where all values are given in the ECI frame, $\rho = \sqrt{(x - x_{si})^2 + (y - y_{si})^2 + (z - z_{si})^2}$ is the range, and the si subscript denotes a ground station coordinate. Values for measurement noise and the

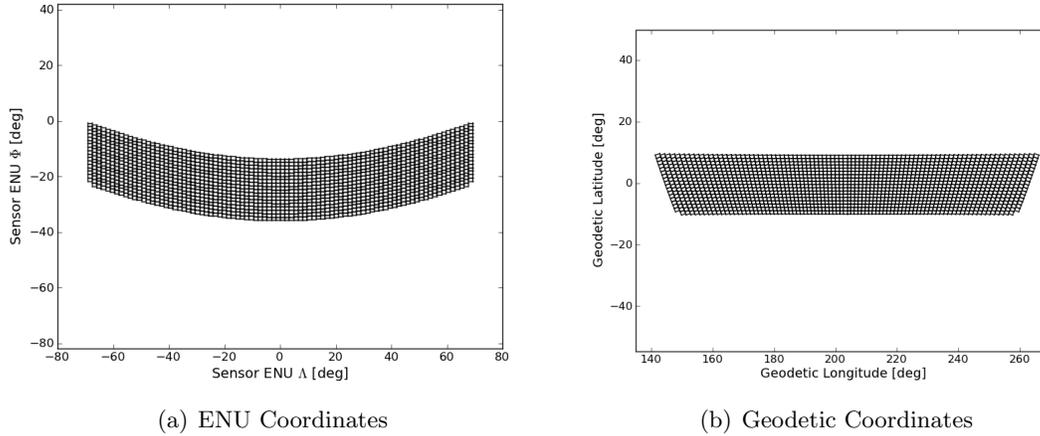


Figure 3.5: Maui Reduced FOR Grid

sensor FOV are provided in Table 3.2 [21]. The offset in hours from GMT is provided for the date February 27, 2013 used as the initial epoch in the simulation.

Table 3.1: Sensor Parameters [79]

Site	Latitude [deg]	Longitude [deg]	Altitude [m]	Az Limits [deg]	El Limits [deg]
Socorro, NM	33.82	-106.66	1510.2	[0,360]	[20,90]
Maui, HI	20.71	-156.26	3058.2	[0,360]	[20,90]
Diego Garcia	-7.41	72.45	-61.2	[0,360]	[20,90]

Table 3.2: GEODSS Sensor Parameters [21]

Site	FOV Size [α, δ]	Noise [α, δ]	GMT Offset [hours]
Socorro, NM	[1.61°, 1.23°]	[2'', 2'']	-7
Maui, HI	[1.61°, 1.23°]	[2'', 2'']	-10
Diego Garcia	[1.61°, 1.23°]	[2'', 2'']	+6

3.1.3.2 Sensor Tasking Modes

This chapter considers several sensor tasking modes which can be grouped into two categories based on whether they use a top-down information theoretic approach or bottom-up composition of heuristic rules. The first bottom-up approach considered is a box scan, in which the sensor is tasked to move through the reduced FOR grid one bin at a time until the entire grid is covered, and

then the pattern is repeated. The baseline version of this scan schedules observations at 1 minute intervals; however, the sensor only requires 6 seconds to generate an observation [21], so a faster box scan is also employed. The fast box scan operates at a 10 second interval to allow for slew rate and settling time, and is able to cover the reduced FOR about 2.5 times in a 12 hour tasking window.

The information theoretic approaches define pointing tasks based on known object locations, allowing for a continuous range of pointing angles within the FOR. They generally favor going after opportunities to observe multiple objects, as well as objects with large uncertainties in the in-track and cross-track directions where angles-only measurements provide the most benefit. The CPHD Rényi divergence is used as the reward function, computed from the estimated multitarget state within the field of view only. Therefore, the summation in Eq. (3.18) is taken over the components with means in the FOV, not the full PHD. The cardinality $p_0(n)$ in the FOV is computed as a multi-Bernoulli distribution [49], $p_0(n) = \left(\prod_{j=1}^{J_{\text{FOV}}} (1 - q_j) \right) \cdot \sigma_n \left(\left\{ \frac{q_1}{1 - q_1}, \dots, \frac{q_{J_{\text{FOV}}}}{1 - q_{J_{\text{FOV}}}} \right\} \right)$, where $\sigma_n(\cdot)$ are the elementary symmetric functions, and the component weights are used for the values of q_j , setting a maximum value $q_j = 0.999$ for any case $w_j \geq 1$.

Figure 3.6 provides a conceptual diagram of the sensor tasking and filtering process. Beginning with an estimated multitarget state \hat{X}_k at time t_k , the sensor allocation algorithm computes a predicted state at the next time step \bar{X}_i . Hypothetical pointing tasks are determined for each known target in the field of regard, and the expected information gain in the FOV is computed and stored in an assignment table. The process is repeated for each time step in the assignment window, after which the table is processed by the auction algorithm [5] to determine a set of sensor tasks. The simplest approach is a single-step algorithm, in which the length of the assignment window $T_w = 1$, and the auction assignment outputs the target ID associated with the largest expected information gain. In a multistep algorithm, the auction assignment is used to maximize the information gain over a window with $T_w > 1$, as described in detail below. With the sensor tasks selected, measurements are simulated and processed by the filter prior to computing the expected information gain for the next assignment window.

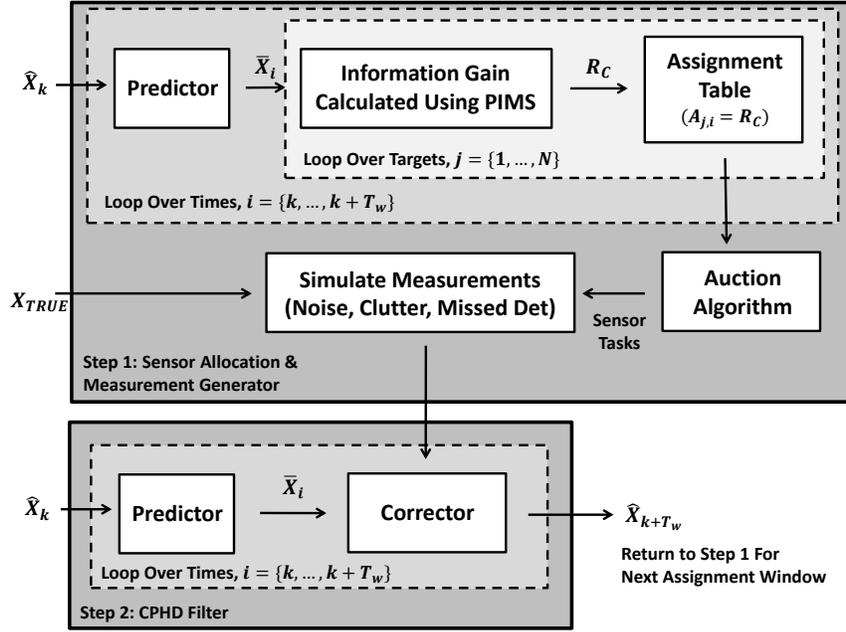


Figure 3.6: Sensor Tasking and Filter Block Diagram

The multistep information theoretic sensor allocation scheme is designed to maximize information gain over a given observation window without repeating tasks. If an object is expected to drift out of the field of regard, the sensor will attempt to collect a measurement while the opportunity still exists, which is not a capability of the single-step algorithm. The ability to account for objects leaving the field of regard is handled automatically by the auction algorithm, described in detail below, and does not require any additional considerations or target priorities to be set as part of the tasking reward function. Generally multistep sensor management is expected to outperform single-step methods, provided the system dynamics are well understood and target states and uncertainties can be reliably predicted throughout the observation window [50]. This presents an increased challenge for GEO SSA, due to the nonlinear dynamics which cause the predicted uncertainties to become non-Gaussian over time. As the focus of this chapter is the application of various sensor allocation schemes, the analysis presented is confined to the use of Gaussian uncertainties propagated using the unscented transform, and does not make use of the AEGIS algorithm. While noting that this may degrade the overall system performance as compared to more princi-

pled nonlinear uncertainty propagation techniques, results in Section 3.2.3 indicate the information theoretic approaches, and in particular the multistep methods, perform well in terms of the number of measurements collected and the average estimated state error.

Several variations of the multistep algorithm are implemented to maximize information gain over observation windows of different lengths. The problem is formulated as a 2D assignment as shown in Fig. 3.7, where each row represents a pointing task defined by the location of a known target and each column is a time step in the observation window. The score for each assignment is the information gain, with a zero representing an object that is not in the field of regard. Within the context of the auction algorithm, this indicates an assignment that is not allowed. The algorithm runs until each column is assigned exactly once, without repeating any rows, such that the total score from all assignments is the highest possible. After the assignment is completed, measurements are simulated and processed for the whole observation window, then the process is repeated as depicted in Figure 3.6.

Target ID	t_0	t_1	t_2
1	⑧	9	10
2	0	0	0
3	0	12	⑬
4	11	⑭	0

Figure 3.7: Sample Multistep Assignment

Figure 3.7 provides a simple example for 4 objects and 3 time steps. At the first time, target 1 is selected, even though it is the second highest score at the time. This assignment allows object 4 to be selected at the second time, when the expected information gain is high. The multistep approach also ensures object 4 is observed before it leaves the field of regard at the final time. Object 2 is not selected at any time as a result of not being in view.

Three different assignment window lengths are considered, corresponding to the value of T_w

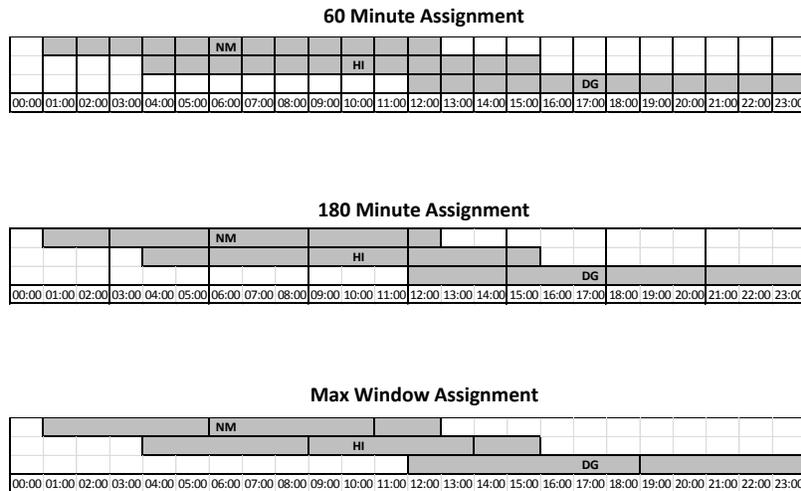


Figure 3.8: Multistep Assignment Windows

in Fig. 3.6. Each sensor is modeled as active for 12 hours per day, from local 6 PM to 6 AM, denoted by the gray boxes in Fig. 3.8 with times given in GMT. The three sensors from Table 3.1, Socorro (NM), Maui (HI), and Diego Garcia (DG) are all shown. The first case considered is a 60 minute assignment. At the end of each hour, measurements are simulated for each active sensor and processed by the filter to produce an updated multitarget state estimate for use in the next assignment block. The 180 minute assignment is similar, with the exception that in certain cases a shorter 1 or 2 hour assignment window is used when a sensor transitions between active and inactive. Measurements are processed by the filter in 3 hour blocks as shown in the figure, and the shorter windows ensure that each assignment block for each sensor uses the same estimated multitarget state, i.e., the input state \hat{X}_k in Figure 3.6 is synchronized between sensors and represents the latest available estimate. The length of time between processing measurements in the filter may also be referred to as the filter update interval.

The final assignment window considered is a variable length maximal assignment. The auction algorithm requires at least as many non-zero rows as there are columns to compute an assignment.

In this application, a maximal assignment corresponds to using exactly as many time steps as there are objects visible to the sensor at some point in the observation window. For example, if 300 objects are visible from Socorro at some point over the course of 300 minutes, then a 5 hour assignment is computed. Because the number of objects visible to each sensor is different and changes throughout the day, the length of the assignment window is variable, and the completion time of each assignment is not synchronized between sensors, as shown nominally in the figure. For this reason, measurements are simulated and stored at the end of each assignment block, but the filter is only used to process measurements at the end of a 24 hour day. This ensures each sensor assignment is computed using the same estimated multitarget state, but prevents the algorithm from using the latest available information on objects and uncertainties when computing the second and third assignment blocks for a given sensor.

By its construction, the maximal assignment is designed to schedule all visible objects at least once, as the auction algorithm will select each object at exactly one time step in the observation window. In the case that the probability of detection is high and all objects are visible to the sensor network at some point in time, this corresponds to a high likelihood of observing every known object. Over a non-maximal assignment window, there is no guarantee that an object will be scheduled prior to leaving the field of regard. As a point of comparison, a final sensor allocation scheme is considered in which a maximal assignment is computed without using information gain, simply using a fixed value as the score for all objects visible. The assignment windows are constructed in exactly the same manner, but a score of 1 is used for all objects at any time they are visible. This means the observations are collected in an arbitrary order, with no consideration for the best possible conditions to observe each object, while still ensuring that each is scheduled at some point in time. A summary of all the tasking modes is provided in Table 3.3.

Table 3.3: Tasking Modes

Tasking Scheme	Observation Interval	Reward Function	Filter Update Interval
Box Scan	1 Minute	N/A	1 Minute
Fast Box Scan	10 Seconds	N/A	10 Seconds
Max Assign Even	1 Minute	1 if in FOR	24 Hours
Single-Step IG	1 Minute	R_C in the FOV	1 Minute
60 Minute IG	1 Minute	R_C in the FOV	60 Minutes
180 Minute IG	1 Minute	R_C in the FOV	180 Minutes
Max Assign IG	1 Minute	R_C in the FOV	24 Hours

3.2 Numerical Simulation

3.2.1 Evolution Of Information Gain In Time

To demonstrate the behavior of the information gain calculation, a small sample of five objects from the public TLE catalog are simulated over a one day period. Two objects are close together and visible at the same time in the FOV, and the other three are separated such that only one target is visible for a given sensor task, as illustrated in Fig. 3.9(a). Of the lone objects, one is initialized with a large radial uncertainty, one with a large in-track uncertainty, and the final with a large cross-track uncertainty. The mean and covariance matrix for a given object's state are propagated using the unscented transform and information gain computed at 10 minute intervals using the PIMS approximation, though no actual measurement update is performed by the filter. As noted above, the assumption of Gaussian uncertainty propagation is maintained and the results are reflective of changes in information gain when assuming a Gaussian posterior distribution.

Table 3.4: Initial State Uncertainties

Parameter	Value
σ_a	1.0 km
σ_e	10^{-4}
σ_i	$(10^{-2})^\circ$
σ_Ω	$(10^{-4})^\circ$
σ_ω	$(10^{-4})^\circ$
σ_M	$(10^{-2})^\circ$

To incorporate the increased uncertainties, the initial covariance is generated in two steps.

First, a diagonal covariance matrix is created in orbital element space, using the standard deviations in Table 3.4. This covariance is converted to Cartesian ECI using the unscented transform [85]. A small value is chosen for the parameter α to ensure no eccentricity values below zero are computed as part of the transform. A diagonal covariance matrix is created in the Radial/In-track/Cross-track (RIC) frame, with a standard deviation of 10 km in the prescribed direction and zero elsewhere. This covariance is rotated to the ECI frame and added to the original to produce the initial covariance for the simulation. No additional uncertainty is added for the velocity states. The two close objects are initialized with the covariance from the orbital elements only.

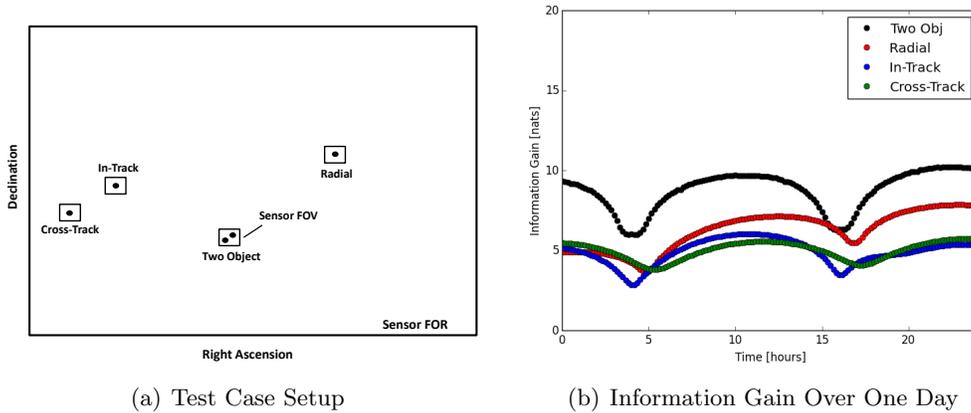


Figure 3.9: Evolution of Information Gain

As seen in Fig. 3.9(b), the information gain for the two close objects is the largest throughout the day. At the initial time, the radial case yields the lowest expected information gain because the angles-only measurements provide little information in the radial direction as a result of the line-of-sight vector and radial position component being approximately parallel. The value increases through the day as the large initial radial uncertainty produces increasing uncertainty in the in-track direction. The in-track and cross-track cases are largely governed by a periodic effect, and show only a modest increase as a result of the initial uncertainty in semi-major axis, which also translates to increased in-track uncertainty over time. In all cases, the expected information gain varies significantly over time, and the figure emphasizes the importance of its use as a reward

function to intelligently select when to view each object. By observing objects at times when the covariance is oriented to take the most advantage of angles measurements, the information theoretic sensor allocation schemes provide a reasoned approach to efficiently collect data on the multitarget state, which is expected to improve the estimation accuracy as compared to heuristic approaches.

3.2.2 Test Case Description

To verify the performance of the sensor allocation schemes described in Section 3.1.3.2, a large scale simulation is included, in which tracking is scheduled for a total of 940 objects in near-geosynchronous orbits over the course of four days using three GEODSS-like sensors. The objects are selected from the public TLE catalog for the date February 27, 2013, and constrained by the following criteria:

$$0 \leq e \leq 0.3 \quad 0^\circ \leq i \leq 70^\circ \quad 0.9 \leq n_m \leq 1.1 \left[\frac{\text{rev}}{\text{sidereal day}} \right]$$

The variations allowed in these orbital elements cause many of the objects to drift relative to the FOR. The objects are additionally constrained to be visible from the box scan FOR at some point in time during the simulation, so that if the box scan fails to detect certain objects, it is only due to the failure of the scheme and not because the objects are not available. Of the 940 objects, 81 are visible in the box scan FOR throughout the course of the entire simulation, with the remaining 859 visible only at some times. Fig. 3.10 provides a visualization of the object catalog, showing the variations in the radius of periapsis and apoapsis, as well as eccentricity and inclination.

Figure 3.11 provides a visual representation of each ground station and its full field of regard, as well as the reduced FOR used in the box scan tasking mode. The three ground stations are capable of near global coverage with a 20° elevation mask applied, with the exception of a gap over the Atlantic Ocean. Note that due to its latitude, the Socorro FOR extends over the north pole, and that there is a large overlap in the coverage from Socorro and Maui. Object locations at the epoch time are also plotted, from which it is clear that most objects are located near the equator.

The filter is initialized with a PHD function of 940 components with the initial uncertainty

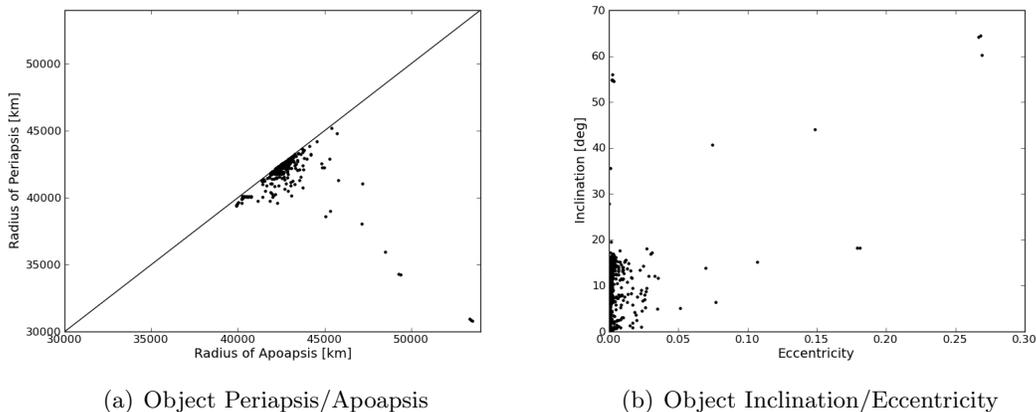


Figure 3.10: Object Catalog Characterization

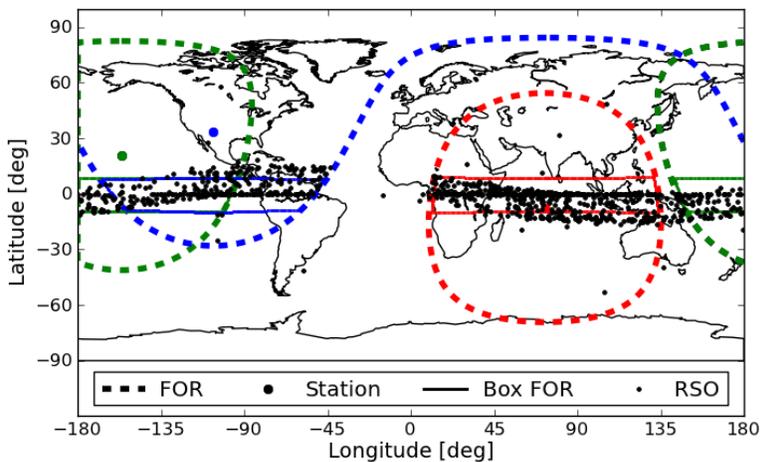


Figure 3.11: Ground Station Coverage and Object Locations at Epoch

for each specified in orbital elements in Table 3.5. The mean estimated state of each object is randomly perturbed from the truth using these values, then both the mean and covariance are converted to Cartesian ECI coordinates using the unscented transform as before [85]. The objects are propagated assuming two-body dynamics, with perturbing forces modeled by a 2x2 spherical harmonics gravity field based on the EGM2008 model [23], solar radiation pressure (SRP), and

luni-solar perturbations using the JPL design ephemeris 430 [58, 10]. The SRP force is modeled assuming all objects are spherical with area-to-mass ratio 0.05 kg/m^2 and reflectivity $C_r = 1.5$. No unmodeled accelerations are included in the filter. As before, the covariance is assumed Gaussian and propagated using the unscented transform. A simple process noise model is used, in which a diagonal process noise covariance matrix $Q = Q_0\Delta t$ is added to the predicted covariance at each time. The matrix Q_0 uses standard deviations of 10^{-4} km and 10^{-7} km/s for each of the positions and velocities in ECI, and Δt is the time interval.

Table 3.5: Initial State Uncertainties

Parameter	Value
σ_a	1.0 km
σ_e	10^{-4}
σ_i	0.01°
σ_Ω	0.01°
σ_ω	0.01°
σ_M	0.01°

No target birth or death is considered, and the simulation does not model the sensor slew rate or settling time, local weather conditions, or lighting conditions of the objects being observed. The filter does model a 0.95 probability of detection for objects in the field of view and assumes a Poisson clutter model with a mean rate of 20 returns each epoch uniformly distributed in the FOV. The filter uses merging and pruning thresholds $U = 9$ and $T = 0.001$ to maintain the Gaussian mixture approximation [80].

3.2.3 Simulation Results

Figure 3.12 provides the average 3D position error over all 940 objects for each of the sensor allocation schemes considered. The position error is computed by taking the maximum weighted GMM component mean for each object and computing the Euclidean distance between the estimate and its true location. Velocity errors are similarly computed and observed to follow the same trends as position errors, and are therefore omitted. From the figure, it is clear that the box scan performs poorly, and it is the only case that does not produce a reduction in the average position error. By

contrast, the remaining sensor allocation schemes converge on average position errors under 2 km by the final time, shown in the close up view in Fig. 3.12(b). The information theoretic schemes all outperform the fast box scan and evenly weighted maximal assignment, confirming the assertion that the information gain reward function drives tasking to collect measurements efficiently to reduce state errors. The best performance is achieved by the maximal information gain assignment scheme, which produces a final multitarget estimate with average position errors of several hundred meters, which is on the order of the prescribed measurement noise.

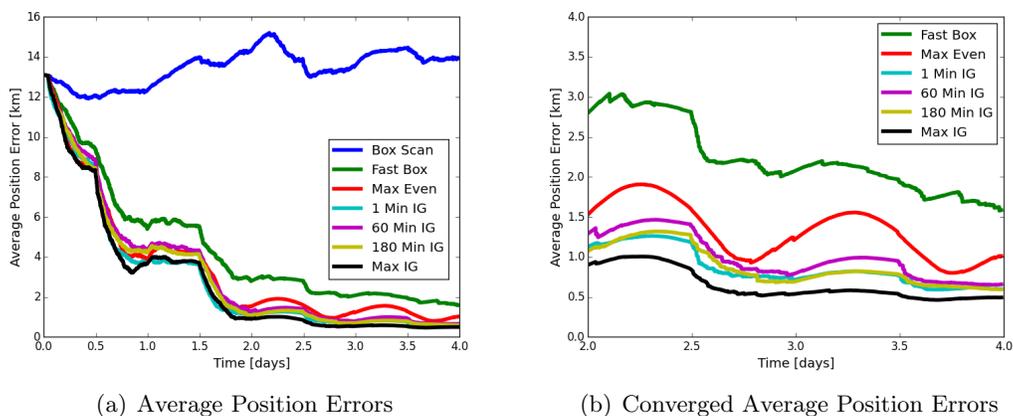


Figure 3.12: Simulation Position Error Results

Analysis of the number of measurements and objects detected by each scheme provides insight into the varying levels of state estimate accuracy, as depicted in Fig. 3.13 and summarized in Table 3.6. The box scan produces the fewest measurements, and fails to detect a large portion of the 940 object catalog due to the lack of an objective function driving the sensor to collect measurements at every epoch. Instead, the sensor spends a large amount of time looking at empty space and only gathers useful data at arbitrary intervals. By increasing the rate of measurement collections, the fast box scan produces approximately six times as many measurements and detects all but 14 objects, yielding a significant improvement in the average position error.

The remaining schemes are all driven by objective functions to ensure they never task the sensor to observe empty space, and all produce correspondingly higher numbers of measurements

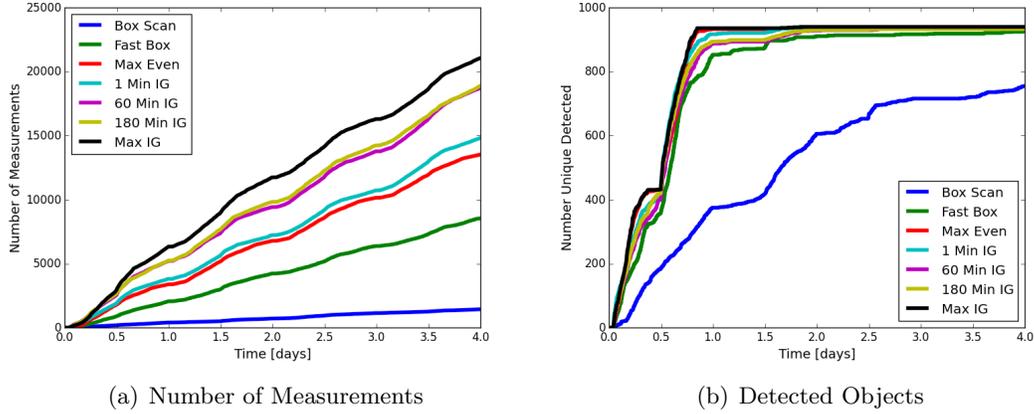


Figure 3.13: Number of Measurements and Objects Detected

Table 3.6: Results Summary

Tasking Scheme	Number of Objects Missed	Number of Observations
Box Scan	185	1441
Fast Box Scan	14	8537
Max Assign Even	1	13512
Single-Step IG	3	14786
60 Minute IG	6	18689
180 Minute IG	7	18842
Max Assign IG	0	21032

and fewer missed objects than either box scan scheme. The information theoretic schemes produce increasing numbers of measurements as the length of the assignment window increases, with the maximal assignment producing the most measurements. In addition, it is the only method considered that detects all objects in the catalog. The evenly weighted maximal assignment scheme also performs well, though it does miss detecting one object. The missed object is a GPS augmentation satellite over the Pacific Ocean, visible from Maui for less than one hour at the end of each night due to the constraints applied in the simulation. By the construction of the tasking windows, this means it is never included in a maximal assignment window and there is no guarantee of its being scheduled. To address the issue, the tasking windows could be constructed differently, or the evenly weighted maximal assignment could be augmented by an additional rule to ensure that it attempts to observe objects that have not been detected before considering other assignments. Note that

the maximal information gain assignment handles this issue automatically. With other objects visible from Maui having been observed, the expected information gain for the undetected object is relatively high, causing the scheme to select the object.

While the evenly weighted maximal assignment yields the best average position error of the non-information theoretic approaches, it still varies significantly over the last two days of the simulation, as opposed to the maximal information gain assignment which has converged to a steady value of several hundred meters. This behavior is partially explained by the fact that the information gain method produces nearly twice as many measurements. It is also due to maximizing the effect of measurements by scheduling opportunities to produce the greatest reduction in uncertainty, an effect that is demonstrated in Fig. 3.14. The plots depict the position errors of all objects in the catalog, with the average values indicated. On the left, the evenly weighted maximal assignment shows a reduction in the number of objects with large errors, but over the last two days, many are still varying in the tens of kilometers, producing a fluctuating average error greater than 1 km. On the right, the maximal information gain assignment shows only a handful of objects varying at this level, particularly by the 3 day mark, after which very few have errors greater than 5 km. The result is a steady low average position error over the final two days, even though the largest individual error is higher than that of the evenly weighted scheme.

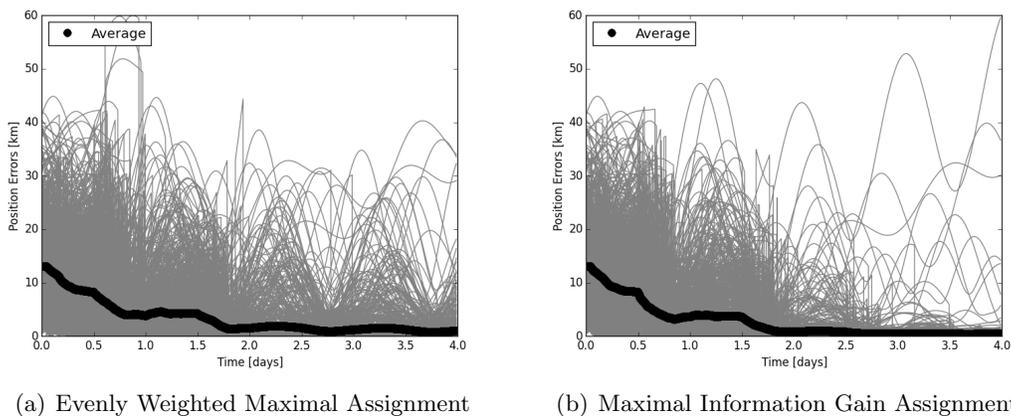


Figure 3.14: Individual Object Errors and Average

The information theoretic tasking schemes all produce similar average position errors and all outperform the non-information theoretic schemes in both accuracy and number of measurements collected. The biggest distinction between them is the number of objects missed in the course of the simulation. The single-step method misses three objects, all of which are super-synchronous and drift into the region that is not covered by any of the sensors. Two of the objects are not scheduled at any point, a result of their having low expected information gain in the time they are available in the field of regard. The third object is scheduled but missed due to the non-unity probability of detection, and it is not rescheduled before drifting out of the FOR. The two objects not scheduled demonstrate an important difference between the single-step and maximal assignment, which has the capacity to anticipate future outages for objects and schedule them while they are still available. The object missed detection is not something that multistep assignment approaches are better suited to handle, it is simply a circumstance of the test. It should be noted that if an additional sensor were used to ensure full global coverage, or if the duration of the test were extended, the single-step method would be expected to detect all objects.

Similarly, the 60 and 180 minute information gain assignments miss several objects. Both schemes fail to schedule the same two objects as the single-step method. This results from the expected information gain being outside the highest range of values for the assignment window length, i.e., if an object is available for less than 60 minutes and the expected information gain is not in the 60 highest values, it will not be scheduled in the 60 minute assignment. The other objects are scheduled but missed detections; they are all super- or sub-synchronous and drift into the region of no coverage before they can be rescheduled. The fact that more missed objects occur for the multistep cases is likely due to the length of the assignment window, which prevents quick rescheduling. While the single-step method can reschedule a missed object immediately, the multistep methods must wait until the entire assignment window is complete before computing a new set of tasks. There is no guarantee an object will be rescheduled before leaving the field of regard for any scheme, and as before, with full global coverage or a longer simulation, all information gain schemes would be expected to detect all objects. For these reasons, the effects of missed objects

should not be weighed too heavily against a scheme’s performance, and a final analysis is considered in which the position errors have been averaged over only the objects detected for each case.

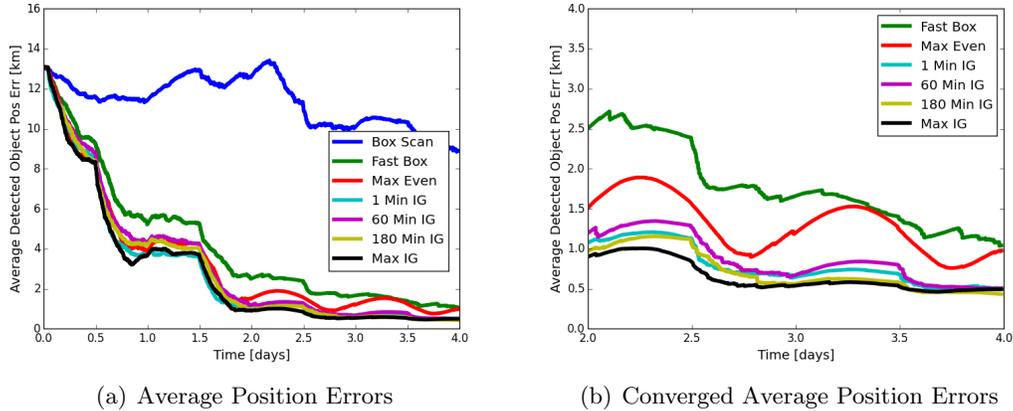


Figure 3.15: Position Error Results of Detected Objects

Figure 3.15 provides the average position errors for each scheme using only the objects that are detected in the course of the simulation. While the box scan is still the worst case, the average fast box scan position error matches that of the evenly weighted maximal assignment by the final time, an indication that for the objects they detect, the two approaches achieve similar performance. The maximal information gain assignment still outperforms the evenly weighted maximal assignment, meaning the differences observed in Fig. 3.12 are not caused by the missed object but rather by the reduction of the majority of individual object errors depicted in Fig. 3.14. Most significantly, the various information theoretic schemes all converge to the same error level. Given a set of sensors to provide full global coverage, any of these schemes would be expected to detect all objects and thereby produce equivalent accuracy in state estimation. The final difference between the information theoretic and non-information theoretic average position errors is several hundred meters, and it is clear that to achieve the best possible accuracy, the information theoretic approaches provide an advantage.

3.3 Chapter Summary

The theory and application of information theoretic sensor tasking has been presented. Through simple demonstration, justification has been provided for the selection of the Rényi divergence, a measure of relative information gain, as a more appropriate reward function for GEO SSA than the Cauchy-Schwarz divergence. Several sensor tasking schemes of varying complexity were proposed and tested in a large scale simulation. The results indicate that a small number of sensors is sufficient to maintain an object catalog if tasked efficiently, and that information theoretic schemes are well suited to this application. While less efficient at collecting measurements, the box scan may be useful as a search mode for new objects, and provided tasks are scheduled at a sufficient rate, it can produce a large number of measurements to help maintain the existing catalog. The box scan tasking mode also poses less computational burden due to its simplicity and the fact that it does not require computing a numerical integral. With full coverage of the geosynchronous belt, either single-step or multistep information gain methods are expected to produce similar performance. It should be noted that the schemes considered can all be used to generate an advance schedule for future time windows, but if real-time or near real-time filtering and sensor allocation is possible, the single-step method offers an advantage in being able to retask sensors quickly. This capability is utilized in the next chapter, in conjunction with the box scan search mode, to detect and schedule follow-on tracking for newly discovered objects.

Chapter 4

Initial Orbit Determination and Follow-On Tracking

Another key problem in SSA research is generating tracks for newly created or discovered objects. Within the context of the CPHD filter, different methods exist to initiate new tracks. The classical derivation of the CPHD filter employs a birth model to account for the appearance of new targets [48], which requires knowledge of the birth process PHD and cardinality distribution as part of the prediction step. Early implementations such as those of Vo et al. [83] assume a Poisson number of new targets originating from predetermined locations designed to cover the entire single target state space. More recent work has relaxed this assumption, determining the birth PHD and cardinality based on measurements, with distributions in state space approximated as Gaussian [65] or partially uniform [2]. Measurement-based birth models offer the advantage of ensuring sufficient density in the PHD function at likely locations of new targets without having to cover the entire state space, which for the orbital debris problem is prohibitively expansive. An alternate approach to reduce the region considered for new targets is to augment the CPHD filter with a spawning model to initiate new tracks in the vicinity of known targets [8]. The spawning model simplifies the track initiation process by taking advantage of knowledge of existing targets; however, it is a less general approach, and offers no means to account for new objects due to space vehicle launches. A simplified version of the CPHD spawning model is presented in Appendix E, but its use is not considered further here. Instead, this chapter seeks to apply a measurement-based birth model to initiate new tracks in the CPHD filter, and to apply the previously developed information theoretic sensor tasking scheme to schedule follow-on measurements for newly discovered objects.

Application of the measurement-based birth model for tracking geosynchronous space objects requires a solution to the initial orbit determination (IOD) problem for angles-only measurements. Classical approaches such as double-r iteration [79] and the methods of Gauss, Laplace, and Gooding [79, 20, 29] rely on having a set of angular measurements from at least three different times, providing at least six independent values to produce a six parameter initial orbit description. However, these approaches are not well suited for application to the measurement-based birth model of the CPHD filter, which is designed to instantiate new tracks based on measurements from a single point in time. They also fail to account for measurement ambiguity due to clutter or misidentified targets.

More recent work considers solutions to the too short arc problem, performing a regression on a single arc of angular measurements to produce a 4D set of angles and rates at one point in time [51]. The 4D measurement set can be used to define an admissible region (AR) in range/range-rate space corresponding to objects captured in Earth orbit [52, 76]. Admissible regions generated from two separate measurement arcs can be correlated to determine if they represent the same object, and if so, produce an initial orbit estimate [51, 26]. In the context of multitarget filtering, both the classical and AR correlation methods pose the challenge of maintaining knowledge of measurement-to-track associations through time, in effect requiring a solution to the multiscan assignment problem [60]. Alternately, the admissible region from a single arc can be approximated using a GMM and used to initiate a filter, which refines the estimate based on subsequent measurements [18]. This latter approach fits the needs of the CPHD filter birth model, and its use has been demonstrated in simulated multitarget GEO SSA scenarios involving clutter and missed detections [35, 73].

This chapter extends previous work by incorporating a sensor allocation scheme to produce the follow-on measurements necessary to refine the initial AR solution, and by implementing a track confirmation process to preclude spurious tracks generated by clutter measurements. In conjunction with the IOD and filtering steps, the scheme is responsible for searching for new objects, and for scheduling measurements to improve and maintain knowledge of new and existing catalog objects, a set of tasks collectively known as the search-detect-track (SDT) problem. To accomplish these

objectives, this research combines the box scan search mode and single-step information theoretic tasking described in Chapter 3. The scheme is designed for applications in which a single sensor is available to perform both the search and track functions, making use of a switch point to transition between tasking modes. This ensures the sensor is able to scan through the field of regard at regular intervals to find new targets while also maintaining sufficient knowledge of objects in the catalog to allow for their reacquisition following gaps in sensor availability.

The remainder of this chapter is organized as follows. Section 4.1 summarizes the constrained admissible region (CAR) IOD technique and provides an example for a geosynchronous object. Section 4.2 provides details on incorporating new targets in the CPHD filter using the CAR and track confirmation process. Section 4.3 reviews previously introduced concepts for information theoretic sensor tasking and describes the overall approach to the search-detect-track problem. Section 4.4 provides results using simulated data to demonstrate the efficacy of the method.

4.1 Constrained Admissible Region

The admissible region approach introduced by Milani et al. [52] is a means to limit the possible initial state of a space object to a subset of six dimensional space by using the information available from a single arc of measurement data. The approach presented here uses measurements of topocentric right ascension and declination and their associated rates to define a region in the range/range-rate space. The initial orbit is parameterized using the following state vector:

$$\mathbf{x}_0 = [\rho \dot{\rho} \alpha \dot{\alpha} \delta \dot{\delta}]^T \quad (4.1)$$

In this formulation, realistic limits on the semi-major axis and eccentricity of the orbit can be mapped to limits on range and range-rate to further constrain the solution. To initialize the filter, this region is approximated by a GMM, and mapped to Cartesian ECI coordinates using the unscented transform [85]. The development here follows that presented by DeMars and Jah [19].

The Cartesian position and velocity of a space object can be written as

$$\mathbf{r} = \mathbf{q} + \boldsymbol{\rho} \quad (4.2)$$

$$\dot{\mathbf{r}} = \dot{\mathbf{q}} + \dot{\boldsymbol{\rho}} \quad (4.3)$$

where \mathbf{q} is the ground station position, $\boldsymbol{\rho}$ is the range vector from the ground station to the object, and $\dot{\mathbf{q}}$ and $\dot{\boldsymbol{\rho}}$ are their respective rates, all in the ECI frame. The range and range-rate vectors can be written in terms of unit vectors as follows

$$\boldsymbol{\rho} = \rho \mathbf{u}_\rho \quad (4.4)$$

$$\dot{\boldsymbol{\rho}} = \dot{\rho} \mathbf{u}_\rho + \rho \dot{\alpha} \mathbf{u}_\alpha + \rho \dot{\delta} \mathbf{u}_\delta \quad (4.5)$$

with the unit vectors defined by

$$\mathbf{u}_\rho = \begin{bmatrix} \cos \alpha \cos \delta \\ \sin \alpha \cos \delta \\ \sin \delta \end{bmatrix}, \quad \mathbf{u}_\alpha = \begin{bmatrix} -\sin \alpha \cos \delta \\ \cos \alpha \cos \delta \\ 0 \end{bmatrix}, \quad \mathbf{u}_\delta = \begin{bmatrix} -\cos \alpha \sin \delta \\ -\sin \alpha \sin \delta \\ \cos \delta \end{bmatrix} \quad (4.6)$$

The first constraint on the admissible region is applied using the semi-major axis, or equivalently orbital energy. The two-body orbital energy equation is

$$\mathcal{E} = \frac{\|\dot{\mathbf{r}}\|^2}{2} - \frac{\mu}{\|\mathbf{r}\|} \quad (4.7)$$

The objective is to map the orbital energy and given angles and angle-rates into range/range-rate space. It is first necessary to parameterize orbital energy in terms of the state vector Eq. (4.1).

Define a set of scalar coefficients using the ground station and measurement parameters:

$$\begin{aligned} w_0 &= \|\mathbf{q}\|^2, & w_1 &= 2(\dot{\mathbf{q}} \cdot \mathbf{u}_\rho), & w_2 &= \dot{\alpha}^2 \cos^2 \delta + \dot{\delta}^2, \\ w_3 &= 2\dot{\alpha}(\dot{\mathbf{q}} \cdot \mathbf{u}_\alpha) + 2\dot{\delta}(\dot{\mathbf{q}} \cdot \mathbf{u}_\delta), & w_4 &= \|\dot{\mathbf{q}}\|^2, & w_5 &= 2(\mathbf{q} \cdot \mathbf{u}_\rho) \end{aligned}$$

From Eqs. (4.2)-(4.3), the orbit radius and velocity can then be defined in terms of the unknown range and range-rate.

$$\|\mathbf{r}\|^2 = \rho^2 + w_5 + w_0 \quad (4.8)$$

$$\|\dot{\mathbf{r}}\|^2 = \dot{\rho}^2 + w_1 \dot{\rho} + w_2 \rho^2 + w_3 \rho + w_4 \quad (4.9)$$

From Eq. (4.7), it is possible to substitute and produce a quadratic solution for $\dot{\rho}$ in terms of ρ , \mathcal{E} , and the scalar coefficients.

$$2\mathcal{E} = \|\dot{\mathbf{r}}\|^2 - \frac{2\mu}{\|\mathbf{r}\|} \quad (4.10)$$

$$2\mathcal{E} = \dot{\rho}^2 + w_1\dot{\rho} + F(\rho) \quad (4.11)$$

$$F(\rho) = w_2\rho^2 + w_3\rho + w_4 - \frac{2\mu}{\sqrt{\rho^2 + w_5\rho + w_0}} \quad (4.12)$$

$$\dot{\rho} = -\frac{w_1}{2} \pm \sqrt{\left(\frac{w_1}{2}\right)^2 - F(\rho) + 2\mathcal{E}} \quad (4.13)$$

The result Eq. (4.13) produces the first constraint on the admissible region. Given a value of the orbital energy, it is possible to compute an upper and lower limit on range-rate for any value of range for which the right ascension, declination, and their rates are equal to the values measured. This constraint is easily reformulated in terms of semi-major axis using the identity $\mathcal{E} = -\mu/2a$, and produces a quadratic curve in the range/range-rate space of constant energy (or semi-major axis) for a given 4 parameter measurement set. An example of this curve is shown in Figure 4.1(a), which applies a minimum and maximum constraint in semi-major axis.

The second constraint is placed on eccentricity, derived in similar fashion using the relationship

$$e = \sqrt{1 + \frac{2\mathcal{E}\|\mathbf{h}\|^2}{\mu^2}} \quad (4.14)$$

The eccentricity is therefore a function of the orbital energy and specific angular momentum, $\mathbf{h} = \mathbf{r} \times \dot{\mathbf{r}}$. To parameterize the angular momentum in terms of the state vector, define a new set of vector and scalar coefficients:

$$\mathbf{h}_1 = \mathbf{q} \times \mathbf{u}_\rho, \quad \mathbf{h}_2 = \mathbf{u}_\rho \times (\dot{\alpha}\mathbf{u}_\alpha + \dot{\delta}\mathbf{u}_\delta),$$

$$\mathbf{h}_3 = \mathbf{u}_\rho \times \dot{\mathbf{q}} + \mathbf{q} \times (\dot{\alpha}\mathbf{u}_\alpha + \dot{\delta}\mathbf{u}_\delta), \quad \mathbf{h}_4 = \mathbf{q} \times \dot{\mathbf{q}}$$

$$c_0 = \|\mathbf{h}_1\|^2, \quad c_1 = 2(\mathbf{h}_1 \cdot \mathbf{h}_2), \quad c_2 = 2(\mathbf{h}_1 \cdot \mathbf{h}_3),$$

$$c_3 = 2(\mathbf{h}_1 \cdot \mathbf{h}_4), \quad c_4 = \|\mathbf{h}_2\|^2, \quad c_5 = 2(\mathbf{h}_2 \cdot \mathbf{h}_3),$$

$$c_6 = 2(\mathbf{h}_2 \cdot \mathbf{h}_4) + \|\mathbf{h}_3\|^2, \quad c_7 = 2(\mathbf{h}_3 \cdot \mathbf{h}_4), \quad c_8 = \|\mathbf{h}_4\|^2$$

The angular momentum vector and magnitude can be computed from the definition above, using Eqs. (4.2)-(4.3) and the coefficients.

$$\mathbf{h} = \mathbf{h}_1\dot{\rho} + \mathbf{h}_2\rho^2 + \mathbf{h}_3\rho + \mathbf{h}_4 \quad (4.15)$$

$$\|\mathbf{h}\|^2 = c_0\dot{\rho}^2 + P(\rho)\dot{\rho} + U(\rho) \quad (4.16)$$

$$P(\rho) = c_1\rho^2 + c_2\rho + c_3$$

$$U(\rho) = c_4\rho^4 + c_5\rho^3 + c_6\rho^2 + c_7\rho + c_8$$

The eccentricity Eq. (4.14) can be rearranged, and the quantities $2\mathcal{E}$ and $\|\mathbf{h}\|^2$ replaced with expressions in terms of range and range-rate.

$$2\mathcal{E}\|\mathbf{h}\|^2 = -\mu^2(1 - e^2) \quad (4.17)$$

$$(\dot{\rho}^2 + w_1\dot{\rho} + F(\rho))(c_0\dot{\rho}^2 + P(\rho)\dot{\rho} + U(\rho)) = -\mu^2(1 - e^2) \quad (4.18)$$

The resulting constraint is a quartic equation for range-rate as a function of range for constant values of eccentricity, defined using a final set of scalar coefficients.

$$a_4\dot{\rho}^4 + a_3\dot{\rho}^3 + a_2\dot{\rho}^2 + a_1\dot{\rho} + a_0 = 0 \quad (4.19)$$

$$a_4 = c_0, \quad a_3 = P(\rho) + c_0w_1, \quad a_2 = U(\rho) + c_0F(\rho) + w_1P(\rho),$$

$$a_1 = F(\rho)P(\rho) + w_1U(\rho), \quad a_0 = F(\rho)U(\rho) + \mu^2(1 - e^2)$$

The equation therefore yields 4 roots, with any imaginary solutions discarded, and the solution produces a curve for constant values of eccentricity corresponding to the 4 parameter measurement. The net result of applying both constraints is the definition of a region in the range/range-rate space, the boundaries of which produce an orbit with the limiting values of semi-major axis and/or eccentricity when the given measurement set is used to fill in the state vector.

To illustrate the concept of the constrained admissible region, a simple example is provided using an observation of the Viasat 1 satellite from a ground station in Maui. Simulated angle and angle-rate measurements are computed, and the range/range-rate space is constrained in semi-major axis, $a \in [41764, 42565]$ km, and eccentricity, $e \in [0, 0.1]$. The only available values of range

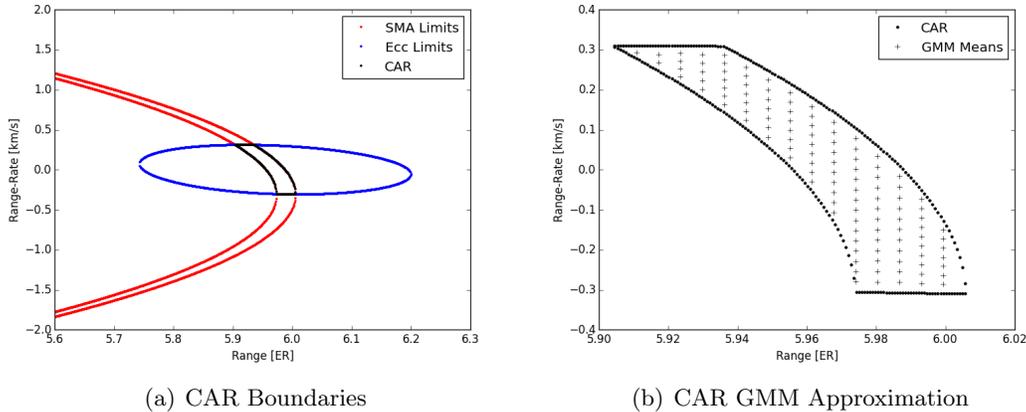


Figure 4.1: Constrained Admissible Region

and range-rate that agree with the measurements and fall within the constraints are those bounded by the black curve in Figure 4.1(a), denoting the CAR.

In the constrained admissible region approach, each point in the CAR is assumed to be equally likely to represent the correct range/range-rate measurement pair. Therefore, the CAR is modeled as a bivariate uniform distribution of possible range and range-rate values. The uniform distribution can be approximated using a GMM, which is then transformed to the ECI frame to produce a suitable input for the CPHD filter. Computing the GMM approximation requires solving a constrained optimization problem with input design parameters to specify the desired variance in range and range-rate, as detailed in Reference [19]. In this example, the design parameters $\sigma_\rho = 5$ km and $\sigma_{\dot{\rho}} = 0.02$ km/s are used, producing a representation of the CAR using a little over 100 GMM components. Figure 4.1(b) provides a close-up view of the CAR with the means of the GMM components indicated.

In this manner, an observation set of four parameters at one time can be used to generate a GMM consisting of many components, at least one of which will be sufficiently close to the truth to produce a good estimate. Subsequent measurements are used to merge and remove components, until a GMM with very few components is left representing the object.

4.2 CPHD Filter Implementation

Having established an approach to initiate tracks based on measurements at a single point in time, it is necessary to develop a method to incorporate new targets in the filter. Without *a priori* information regarding which measurements represent new targets as opposed to existing targets or clutter, all measurements are treated as potential new targets. At time t_{k-1} , each measurement is used to generate a CAR. The combined CAR GMMs then form the basis for the filter birth process used in the prediction step to time t_k . To prevent the accumulation of probability mass as objects are scheduled for repeated observations, this research develops an approach to compute a probability that a given measurement represents a new target based on the estimated target and assumed clutter distributions. The CAR GMM weights for a given measurement are then multiplied by the new target probability, which keeps the sum of new component weights low in the case a measurement is not likely generated by a new target.

This section provides an overview of the measurement-based CPHD filter birth model and a discussion of an augmented version of the filter using a PHD to pre-screen measurements as part of the update. The first topic is development of the new target probability for later inclusion in the birth model.

4.2.1 Computation of New Target Probability

Given a set of measurements and knowledge of the assumed clutter and estimated target distributions, it is possible to compute a probability that a given measurement represents a new target. The proposed method makes use of a ratio of Gaussian likelihoods for existing targets and GMM clutter, and should only be taken as an approximation of the probability that the measurement is generated by a new target. With this caveat, results in Section 4.4 demonstrate that the method functions as desired. The full development of the equations is given in Appendix D, the main results are summarized here. Beginning with the assumption that each measurement represents at most one target, it is possible to define a constraint on the probabilities that measurement

\mathbf{z}_i is generated by an existing target, a new target, or a false alarm.

$$p_{\text{exist}}(\mathbf{z}_i) + p_{\text{new}}(\mathbf{z}_i) + p_{\text{FA}}(\mathbf{z}_i) = 1 \quad (4.20)$$

Assuming the Poisson clutter model described in Chapter 2, the clutter cardinality is given by

$$p_{\kappa}(m) = \frac{(\lambda_{\kappa})^m}{m!} e^{-\lambda_{\kappa}} \quad (4.21)$$

A cardinality distribution can be computed for the existing targets in the field of view using a multi-Bernoulli distribution.

$$p_e(m) = \left(\prod_{j=1}^{J_{\text{FOV}}} (1 - p_D w_j) \right) \cdot \sigma_m \left(\left\{ \frac{p_D w_1}{1 - p_D w_1}, \dots, \frac{p_D w_{J_{\text{FOV}}}}{1 - p_D w_{J_{\text{FOV}}}} \right\} \right) \quad (4.22)$$

The combined cardinality distribution representing clutter and existing targets can then be computed from the convolution of the previous cardinalities.

$$p_{\kappa e}(m) = (p_{\kappa} * p_e)(m) = \sum_{i=0}^m p_{\kappa}(i) p_e(m - i) \quad (4.23)$$

Using the existing target GMM and the assumption that false alarms are uniformly distributed in the FOV, it is possible to compute a conditional probability that \mathbf{z}_i represents an existing target or false alarm, given that m out of m_{k-1} measurements are generated by such,

$$p_{\text{FAE}}(\mathbf{z}_i | \frac{m}{m_{k-1}}) = \frac{\sum_{j=0}^m \frac{p_{\kappa}(j) p_e(m - j)}{\sum_{l=0}^m p_{\kappa}(l) p_e(m - l)}}{\left[\frac{j}{m_{k-1}} + \frac{1}{C_j^{m_{k-1}}} \sum_{\substack{\pi \subseteq Z_{k-1}: \\ |\pi|=j}} p_{\text{exist}}(\mathbf{z}_i | \frac{m-j}{|Z_{k-1} \setminus \pi|}) \right]} \quad (4.24)$$

where $p_{\text{exist}}(\cdot)$ is computed using a ratio of measurement likelihood functions as discussed in Appendix D. With the conditional probability and cardinality distribution, it is possible to compute the probability that \mathbf{z}_i is generated by either an existing target or false alarm.

$$p_{\text{FAE}}(\mathbf{z}_i | Z_{k-1}) = \sum_{m=1}^{m_{k-1}} p_{\kappa e}(m) \cdot p_{\text{FAE}}(\mathbf{z}_i | \frac{m}{m_{k-1}}) \quad (4.25)$$

Finally, the probability that \mathbf{z}_i represents a new target is computed by applying the constraint in Eq. (4.20). The incorporation of this value into the CPHD birth model is discussed next.

$$p_{\text{new}}(\mathbf{z}_i | Z_{k-1}) = 1 - p_{\text{FAE}}(\mathbf{z}_i | Z_{k-1}) \quad (4.26)$$

4.2.2 Measurement-Based Birth Model

This section provides a review of the CPHD filter predictor equations, including the target birth model ignored in previous chapters. The filter corrector is unaffected by inclusion of the birth model, and its discussion is therefore omitted. As before, the PHD is approximated using a GMM.

$$\nu_k(\mathbf{x}) \approx \sum_{j=1}^{J_k} w_k^{(j)} p_g(\mathbf{x}; \mathbf{x}_k^{(j)}, P_k^{(j)}) \quad (4.27)$$

where $w_k^{(j)} > 0$ are the weights and $p_g(\mathbf{x}; \mathbf{x}_k^{(j)}, P_k^{(j)})$ is a multivariate Gaussian PDF with mean $\mathbf{x}_k^{(j)}$ and covariance $P_k^{(j)}$. The prediction step for the PHD and cardinality is given by

$$\nu_{k|k-1}(\mathbf{x}) = \gamma_k(\mathbf{x}) + p_S \sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} p_g(\mathbf{x}; \mathbf{x}_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}) \quad (4.28)$$

$$p_{k|k-1}(n) = \sum_{j=0}^n p_{\Gamma,k}(n-j) \sum_{l=j}^{\infty} C_j^l p_{k-1}(l) p_S^j (1-p_S)^{(l-j)} \quad (4.29)$$

where $\gamma_k(\mathbf{x})$ represents the PHD of birth targets and can be approximated by a GMM as in Eq. (4.27), p_S is the probability of target survival, $p_{\Gamma,k}(\cdot)$ represents the cardinality of target birth and $C_j^l = \frac{l!}{j!(l-j)!}$ is the binomial coefficient.

The computation of new target probabilities lends itself well to modeling target birth as a multi-Bernoulli process. Given the measurement set and associated probabilities for time t_{k-1} , the multi-Bernoulli birth cardinality for time t_k is computed from

$$p_{\Gamma,k}(n) = \prod_{i=1}^{m_{k-1}} (1 - p_{\text{new},i}) \cdot \sigma_n \left(\left\{ \frac{p_{\text{new},1}}{1 - p_{\text{new},1}}, \dots, \frac{p_{\text{new},m_{k-1}}}{1 - p_{\text{new},m_{k-1}}} \right\} \right) \quad (4.30)$$

where the abbreviation $p_{\text{new},i} = p_{\text{new}}(\mathbf{z}_i | Z_{k-1})$ has been adopted. The birth PHD is generated by computing a CAR for each measurement in Z_{k-1} and propagating the mean and covariance to t_k . The weights of the CAR for a given measurement are multiplied by the new target probability, which keeps the added probability mass consistent with the predicted cardinality and minimizes the contribution of new components spawned by measurements that are likely associated to existing targets or clutter.

$$\gamma_k(\mathbf{x}) = \sum_{i=1}^{m_{k-1}} p_{\text{new},i} \sum_{j=1}^{J_{\text{CAR},i}} w_{k-1}^{(j)} p_g(\mathbf{x}; \mathbf{x}_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}) \quad (4.31)$$

Note that the probability of survival does not affect either the predicted cardinality or PHD of birth targets because the CPHD prediction step is predicated on the assumption that birth and target survival are independent processes.

The multi-Bernoulli birth model and CAR IOD technique provide a straightforward means to introduce new targets in the CPHD filter. However, as will be demonstrated in subsequent testing, the method does not produce a reliable estimate of the number of objects. Previous research has noted a tendency for the measurement-based birth model to produce a bias in the cardinality estimate, unless the CPHD measurement update is adjusted [65]; however, the implementation is different from that given here. Ristic et al. instantiate the birth model at the same time measurements are collected, whereas this chapter makes use of measurements from the previous time step to generate new targets at the current time. As a result, the standard CPHD filter update equations are used and a different method is proposed to address the inaccurate cardinality estimate. In this research, the filter is restructured in a two-tier approach, in which new targets are maintained in an unconfirmed PHD mixture until sufficient follow-on measurements warrant their inclusion in the CPHD for existing targets. The approach is similar to other applications of the PHD filter to pre-screen measurements prior to estimation using MHT [56], and produces an estimate of target number robust against the effects of clutter and missed detections.

4.2.3 Augmented CPHD Filter

The methodology of the augmented CPHD approach is provided in Figure 4.2. At time t_{k-1} , a set of m_{k-1} measurements are collected. A measurement update is performed for all objects in the unconfirmed PHD mixture, with the GMM representing confirmed objects incorporated as part of the clutter model. In this manner, measurements generated by confirmed objects are treated as likely false alarms during the unconfirmed PHD measurement update. Subsequently, the measurements are used to update the estimate of objects in the confirmed CPHD, this time taking the GMM of the unconfirmed objects to be clutter. Both update steps use the combined clutter model described in Appendix D to include the standard Poisson false alarms.

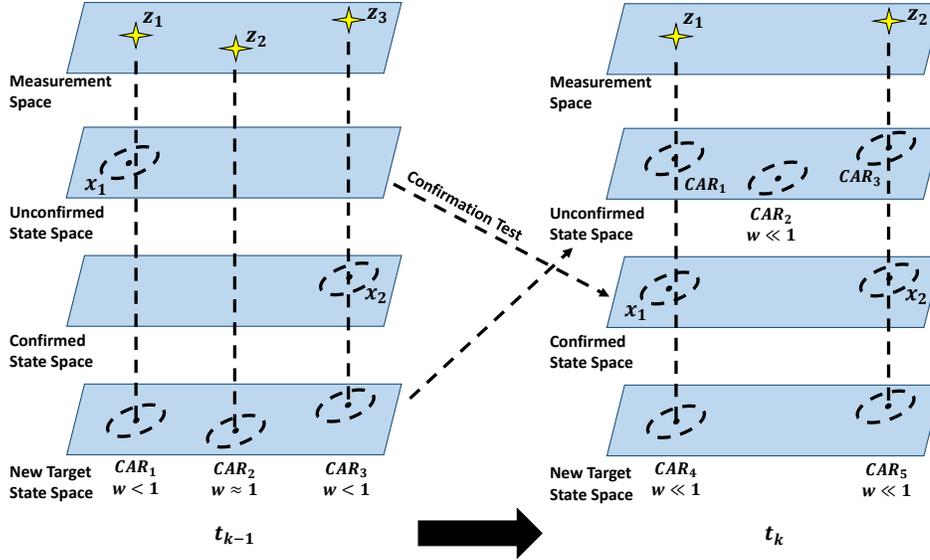


Figure 4.2: Augmented CPHD Filter

As part of the measurement update, a new target probability and CAR GMM are computed for each measurement. In the scenario depicted above, measurement z_2 is not located near any known objects, and spawns a new object with a weight of nearly 1, reduced only by the probability that it is generated as a Poisson false alarm. Measurement z_1 is located near an unconfirmed target x_1 , which results in a reduced probability of representing a new object. Similarly, measurement z_3 is located near a confirmed target. It should be noted that the calculation of $p_{\text{new},i}$ incorporates both the confirmed and unconfirmed multitarget states.

During the prediction step, newly generated targets are moved to the unconfirmed PHD as shown. A test is performed to determine if previously unconfirmed targets should be incorporated in the confirmed target CPHD. When targets are confirmed, the birth cardinality prediction for the confirmed CPHD is formed as a multi-Bernoulli distribution from the weights of the newly confirmed objects.

The following measurement update at time t_k is performed in the same manner. In the case that no measurements are present near newly initialized objects, the original measurement is likely a false alarm and the weights are reduced. Eventually, if no measurements update the estimate,

the components will be removed as part of the filter update. New CAR GMMs are initiated for both measurements in Z_k ; however, since there are both confirmed and unconfirmed tracks near each, the resulting $p_{\text{new},i}$ is low. If the new target probability is below the pruning threshold, it indicates the components are not needed as they would be removed anyway. Computation of the probability therefore prevents unnecessarily adding components to the unconfirmed PHD mixture indefinitely. The unconfirmed PHD filter then serves as an effective screen against clutter, and is also a way to aggregate information from multiple CAR GMMs into refined estimates of targets before being introduced to the main CPHD filter tracking known objects.

4.2.4 CPHD Clustering

In addition to the birth model and PHD pre-filter, the CPHD filter is implemented using a clustering technique to mitigate against an issue known as ‘spooky action at a distance,’ in which missed detections cause probability mass to shift from undetected to detected objects, even if they are physically separated to the point of noninteraction [50, 25]. The feature results from the inherent design of the filter, in which the multitarget state is represented as a density from which likely target states are later extracted, as opposed to strictly enforcing track identities as in MHT. The simulation in Chapter 3 avoids this issue by ignoring target birth and death and maintaining a simple list of track identities; however, this is not sufficient for the newly introduced measurement-based birth model.

The CPHD implementation in this chapter forms clusters using the Mahalanobis distance, such that any GMM components within a threshold U_c of a peak in the intensity function are considered to be part of the same cluster [59]. A CPHD measurement update is performed for each cluster, taking all other clusters and the unconfirmed PHD mixture as a GMM clutter model as described in Appendix D. By performing a measurement update for each cluster separately, weights are not allowed to shift to other clusters, and the spooky action effect is avoided.

4.3 Incorporation of Sensor Allocation Scheme

The final component of the search-detect-track algorithm is the sensor allocation step, which must perform both the search and track functions. The box scan mode discussed in Chapter 3 is used to perform the search, at a time interval of one minute in order to allow computation of the 4D angle and angle-rate measurement vector needed to generate the CAR. To schedule follow-on tracking, the single-step information theoretic approach is used, with the CPHD form of the Rényi divergence as the objective function. The switch point between modes is defined as a threshold in the expected information gain. At each time step, the Rényi divergence is computed for all known objects, from the confirmed and unconfirmed mixtures in the case of the augmented CPHD filter, as described in detail below. If the maximum information gain is below the threshold, the sensor is tasked to observe the next bin in the search grid. If one or more options yield information gain above the threshold, the sensor is tasked to achieve the maximum information gain.

4.3.1 Review of Information Theoretic Sensor Tasking

The expected information gain is computed using the CPHD form of the Rényi divergence evaluated in the FOV, as described in Section 3.1.3.2. Recall that this involves computing a cardinality $p_0(n)$ for GMM components in the FOV, and is therefore applicable for both the confirmed and unconfirmed state spaces modeled in the augmented filter. The cardinality is computed as a multi-Bernoulli distribution [49], $p_0(n) = \left(\prod_{j=1}^{J_{\text{FOV}}} (1 - q_j) \right) \cdot \sigma_n \left(\left\{ \frac{q_1}{1 - q_1}, \dots, \frac{q_{J_{\text{FOV}}}}{1 - q_{J_{\text{FOV}}}} \right\} \right)$, where $\sigma_n(\cdot)$ are the elementary symmetric functions, and the component weights are used for the values of q_j , setting a maximum value $q_j = 0.999$ for any case $w_j \geq 1$.

The Rényi divergence is computed from Eq. (3.18), reprinted here for reference,

$$\nu_0(\mathbf{x}) \approx \sum_{i=1}^{J_0} w_i p_g(\mathbf{x}; \mathbf{m}_i, P_i) \quad \nu_1(\mathbf{x}; \mathbf{u}) \approx \sum_{j=1}^{J_1} w_j p_g(\mathbf{x}; \mathbf{m}_j, P_j) \quad (4.32)$$

$$R_C(\mathbf{u}) \approx -2 \log \sum_{n \geq 0} \left(\frac{p_1(n; \mathbf{u})}{N_1^n} \right)^{1/2} \left(\frac{p_0(n)}{N_0^n} \right)^{1/2} \cdot \left[\int \left(\sum_{i=1}^{J_0} \sum_{j=1}^{J_1} w_i w_j K_{i,j} p_g(\mathbf{x}; \mathbf{m}_{i,j}, P_{i,j}) \right)^{1/2} d\mathbf{x} \right]^n \quad (4.33)$$

$$\begin{aligned} K_{i,j} &= p_g(\mathbf{m}_i; \mathbf{m}_j, P_i + P_j) \\ P_{i,j} &= [P_i^{-1} + P_j^{-1}]^{-1} \\ \mathbf{m}_{i,j} &= P_{i,j} [P_i^{-1} \mathbf{m}_i + P_j^{-1} \mathbf{m}_j] \end{aligned}$$

where the subscripts 0 and 1 refer to the prior and posterior distributions respectively.

At each time, potential tasks are determined from the estimated target number and peaks of the PHD function. In the case of the augmented filter, the number of tasks considered for the unconfirmed PHD is computed from the EAP target number estimate by taking the sum of weights of the GMM. For the CPHD filter, the number of tasks considered is taken as the MAP of the cardinality. In both cases, pointing tasks are evaluated for the N_k highest weighted components of the GMM representing the PHD.

$$\text{EAP: } N_k = \sum_{j=1}^{J_k} w_k^{(j)} \quad (4.34)$$

$$\text{MAP: } N_k = \operatorname{argmax} p_k(n) \quad (4.35)$$

Suppose at time t_k the sum of weights in the unconfirmed PHD is 3, and the MAP of the confirmed cardinality is 1. The SDT algorithm evaluates 4 potential tasks, 3 at the peaks of the unconfirmed PHD and 1 at the peak of the confirmed PHD. The PIMS approximation is applied, simulating a measurement at each of the N_{FOV} highest peaks within the field of view for each potential task. The expected information gain is computed using the combined GMM models representing both the confirmed and unconfirmed state spaces, so if one target is modeled in each, two measurements are simulated and the information gain reflects the combined CPHD measurement update. Following computation of the expected information gain, if no value is above the user-defined information gain cutoff C , the sensor is scheduled to observe the next bin in the search grid. If some of the tasks exceed the cutoff, the maximum value task is scheduled per the

single-step information theoretic scheme. No multistep assignments are considered in the course of this chapter.

With the sensor task determined, measurements are simulated and processed by the filter. The full search-detect-track algorithm steps are depicted in Figure 4.3 and summarized in Table 4.1.

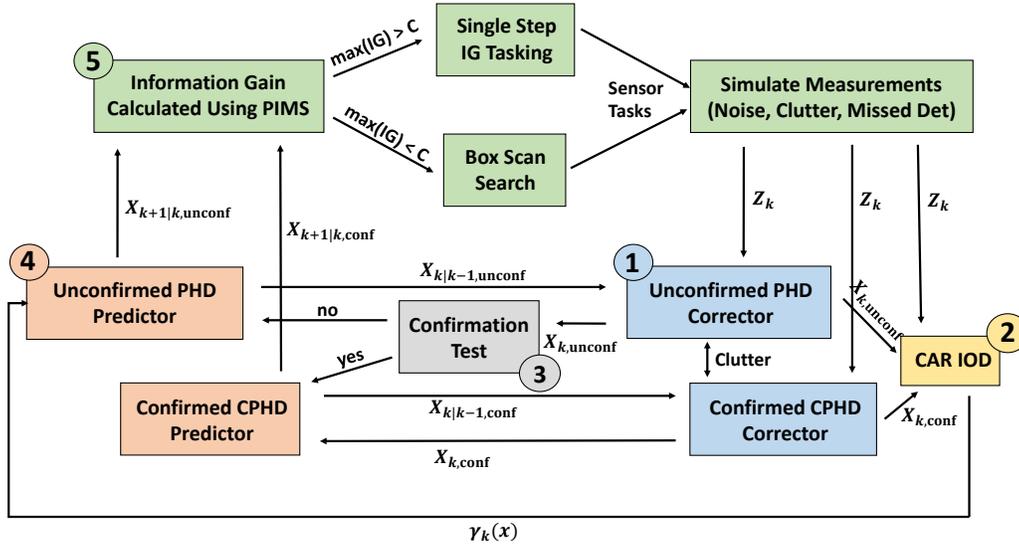


Figure 4.3: Search-Detect-Track Algorithm

Table 4.1: Search-Detect-Track Algorithm

Step 1:	At the current time, compute measurement update for all known objects (confirmed and unconfirmed).
Step 2:	Compute the new target probability for all measurements. Initiate new target estimates using CAR GMM.
Step 3:	Perform test to confirm targets and move to confirmed state space.
Step 4:	Compute the prediction step for confirmed and unconfirmed targets.
Step 5:	Compute the expected information gain for confirmed and unconfirmed targets. <ul style="list-style-type: none"> a. If above threshold, schedule maximum value target for observation. b. If below threshold, task sensor to observe next bin in search grid.
Step 6:	Process measurement update for next time step.

4.4 Numerical Simulation

4.4.1 General Test Conditions and Sensor Parameters

To demonstrate the performance of the search-detect-track algorithm, a series of test cases are developed to track an unknown number of objects from a single GEODSS-like sensor located at Maui, using simulated angle and angle-rate measurements. Much of the test setup is similar to Chapter 3, but the scale has been reduced due to the added complexity of initiating new tracks. The single ground station is modeled using the parameters provided in Table 4.2 [79]. The measurements used are topocentric right ascension and declination and their rates,

$$\alpha = \tan^{-1} \left(\frac{y - y_{si}}{x - x_{si}} \right) \quad \delta = \sin^{-1} \left(\frac{z - z_{si}}{\rho} \right) \quad (4.36)$$

$$\dot{\alpha} = \frac{1}{1 + \left(\frac{y - y_{si}}{x - x_{si}} \right)^2} \left[\frac{(\dot{y} - \dot{y}_{si})(x - x_{si}) - (\dot{x} - \dot{x}_{si})(y - y_{si})}{(x - x_{si})^2} \right] \quad (4.37)$$

$$\dot{\delta} = \frac{1}{\sqrt{1 - \left(\frac{z - z_{si}}{\rho} \right)^2}} \left[\frac{\dot{z} - \dot{z}_{si}}{\rho} - \frac{(\boldsymbol{\rho} \cdot \dot{\boldsymbol{\rho}})(z - z_{si})}{\rho^3} \right] \quad (4.38)$$

where all values are given in the ECI frame and the *si* subscript denotes a ground station coordinate. While the generation of angle-rate measurements is not considered in detail here, the process involves a linear regression over an arc of angle measurements, which produces a reduction in the measurement noise dependent on the length of the arc and the sampling frequency [51]. The values used in this research are consistent with those expected from fitting a dense measurement arc of 20 seconds in length, as proposed elsewhere in the literature [19, 35]. The sensor schedules tasks at 1 minute intervals to allow for slew rate and settling time between generating measurement arcs, though none of these processes are explicitly modeled. At each time, the angles and rates are simulated using Eqs. (4.36)-(4.38) and corrupted by zero mean Gaussian noise with the standard deviations provided in Table 4.2. Poisson clutter is simulated with a mean number of λ_{κ} returns that varies in each test case.

The offset in hours from GMT is provided for the date February 27, 2013 used as the initial

¹ The value given for the altitude here is slightly different from that in Ch. 3. The previous chapter inadvertently uses the location parameters for the Maui MSS sensor instead of the GEODSS. The difference is only 0.4 m and is therefore not expected to influence the results significantly.

Table 4.2: Maui GEODSS Sensor Location and Parameters [79, 21, 35]

Maui, HI	
Latitude [deg]	20.71
Longitude [deg]	-156.26
Altitude [m]	3058.6 ¹
Az Limits [deg]	[0,360]
El Limits [deg]	[20,90]
GMT Offset [hours]	10
FOV Size [α, δ] [deg]	[1.61,1.23]
FOV Size [$\dot{\alpha}, \dot{\delta}$] [$\frac{\text{deg}}{\text{sec}}$]	[0.001, 0.001]
Noise [α, δ] [a-sec]	[0.8, 0.8]
Noise [$\dot{\alpha}, \dot{\delta}$] [$\frac{\text{a-sec}}{\text{sec}}$]	[0.07, 0.07]

epoch in the simulation. The start time for all test cases has been advanced to 04:00 GMT to coincide with 6 PM local time when the sensor is modeled as first being active. As in Chapter 3, the sensor is online for 12 hours followed by 12 hours offline. In order to cover the search space at regular intervals using the 1 minute box scan mode, the effective field of regard is reduced to a range of ± 2 degrees in geodetic latitude, corresponding to approximately 340 bins as shown in Figure 4.4. This allows the box FOR to be scanned twice in the course of a 12 hour tasking session.

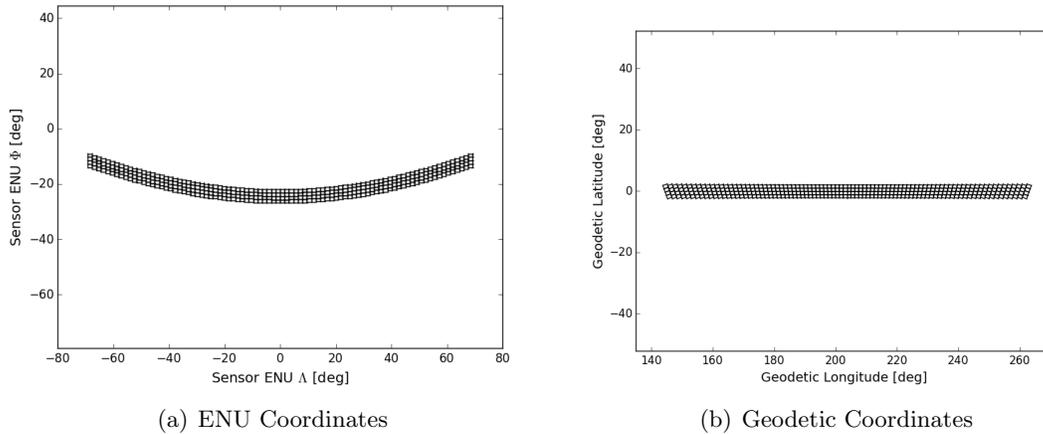


Figure 4.4: Maui Reduced Box FOR

The objects considered in the study are limited to low inclination geostationary orbits to ensure they will be discovered as the sensor steps through the grid and can be accurately represented

by the prescribed CAR IOD parameters. A total of 10 objects are modeled, with the osculating orbital elements and geodetic longitude values provided in Table 4.3 for the initial epoch time. The search grid begins at the easternmost point; objects in the table are listed in the order in which they are expected to be found. Certain test cases model only two objects, as identified in the table. Object states are initialized from the public TLE catalog, transformed to osculating elements.²

Table 4.3: Orbital Elements and Longitude at Epoch

Object	NORAD ID	Longitude	a [km]	e	i	Ω	ω	M
Echostar 10	28935	110.2° W	42165.3	1.55e-4	0.069°	106.2°	-148.6°	89.1°
Mexsat 3	39035	114.8° W	42165.7	3.10e-4	0.059°	109.2°	-73.5°	6.35°
Viasat 1 ³	37843	115.1° W	42165.7	3.02e-4	0.095°	96.2°	-120.5°	65.9°
XM-2	26724	115.2° W	42165.4	3.62e-4	0.011°	88.1°	-128.7°	82.2°
Anik F3 ³	31102	118.7° W	42165.7	2.78e-4	0.108°	99.5°	-113.8°	52.5°
AMC-21	33275	124.9° W	42165.5	2.62e-4	0.034°	98.4°	-113.8°	47.3°
Galaxy 14	28790	125.0° W	42165.9	2.87e-4	0.038°	105.4°	-120.3°	46.8°
AMC-8	26639	139.0° W	42165.8	2.91e-4	0.030°	89.6°	-88.7°	16.9°
AMC-23	28924	172.0° E	42165.5	3.01e-4	0.052°	106.0°	-135.0°	357.9°
Beidou G4	37210	160.0° E	42165.6	3.29e-4	0.520°	10.1°	-140.5°	87.2°

Objects are propagated using the same dynamical model as Chapter 3, which includes the two-body force and perturbations due to a 2x2 spherical harmonics gravity field based on the EGM2008 model [23], solar radiation pressure (SRP), and luni-solar perturbations using the JPL design ephemeris 430 [58, 10]. The SRP force is modeled assuming all objects are spherical with area-to-mass ratio 0.05 kg/m² and reflectivity $C_r = 1.5$. No unmodeled accelerations are included in the filter. As before, the covariance is assumed Gaussian and propagated using the unscented transform. A simple process noise model is used, in which a diagonal process noise covariance matrix $Q = Q_0 \Delta t$ is added to the predicted covariance at each time, using larger values than those of Chapter 3. The matrix Q_0 uses standard deviations of 10^{-3} km and 10^{-6} km/s for each of the positions and velocities in ECI, and Δt is the time interval.

In all cases, the filter is initialized with no prior knowledge of the objects; the cardinality is set to exactly zero targets and the PHD contains no GMM components. New components are created

² TLE obtained from www.space-track.org

³ Object used in 2 object cases

for each measurement using the CAR IOD technique with the parameters described in Section 4.1. Component weights are multiplied by the probability that the measurement represents a new target. In the case of the single thread CPHD filter, the CAR components and birth cardinality are incorporated into the filter birth model for the next prediction step. For the augmented CPHD filter, the GMM components are added to the unconfirmed PHD mixture. Subsequently, a confirmation test is performed for GMM components in the unconfirmed PHD. In this research, a target is confirmed if the GMM representing it contains a component with a weight $w_j > 0.9$. In this case, a cluster is formed for all components within U_c of the new peak and the whole cluster is moved to the confirmed CPHD as part of the birth model. The birth cardinality is modeled as a multi-Bernoulli process using the sum of the weights of the confirmed GMM cluster, for each confirmed target in the case there are multiple confirmations at a given time.

As implemented, the CAR IOD technique produces approximately 100 components for each measurement, which quickly drift apart from one another and become less useful as a probabilistic description of likely object locations. The approach is best suited for prompt follow-on measurements to refine the estimate and confirm targets. Extraneous components in the unconfirmed PHD mixture can negatively impact the validity of the information gain calculation, subsequent sensor tasking decisions, and computation of the new target probability. To mitigate against the accumulation of GMM components that no longer provide useful information on new targets, the unconfirmed PHD mixture is deleted any time the sensor transitions from active to inactive, meaning targets must be confirmed during the same 12-hour observation window in which they are first detected. Any targets dropped before confirmation can simply be detected and confirmed during the next search period.

The GMM mixture is maintained as before by pruning and merging targets, in this case using a pruning threshold $T = 10^{-5}$ and merging threshold $U = 4$. In addition, a maximum number of components is specified, $J_{\max} = 1000$, separately for the confirmed and unconfirmed mixtures. The pruning and merging steps are implemented as described in Appendix B.

4.4.2 Test Case 1: Single Thread CPHD

The first test case considers use of the single thread CPHD filter in conjunction with the SDT sensor tasking algorithm. Two objects, NORAD 37843 and 31102, are modeled over the course of a single 12-hour observation window. Position and velocity errors are computed using the 1st order OSPA of the Euclidean distance between the true and estimated state values corresponding to the N_k highest peaks in the filter. The OSPA metric does not include cardinality errors, it simply represents the average 3D position and velocity errors at each time. The estimated target number N_k is computed as the MAP of the cardinality and reported separately for comparison to the true number of objects detected.

To demonstrate the behavior of the filter birth model, different values of the clustering threshold, mean clutter rate, and probability of detection are considered. Four sets of values are used, with parameters summarized in Table 4.4. In the first case, denoted ST1A, the clustering threshold is set to $U_c = 36$ while no clutter or missed detections are included in the measurements, though the filter is initialized with the values $p_D = 0.99$ and $\lambda_\kappa = 0.01$ to avoid numerical issues. The second test case, ST1B, increases the clustering threshold to $U_c = 100$ while maintaining the same measurement conditions of no clutter or missed detections. Tests ST1C and ST1D set the clustering threshold $U_c = 100$, while adjusting the mean clutter rate and probability of detection to $(\lambda_\kappa = 0.01, p_D = 0.99)$ and $(\lambda_\kappa = 0.05, p_D = 0.95)$ respectively. The information gain switch point is set to $C = 5$ for all cases.

Table 4.4: Test Case 1 Parameters

	ST1A	ST1B	ST1C	ST1D
Measurement λ_κ	0.0	0.0	0.01	0.05
Filter λ_κ	0.01	0.01	0.01	0.05
Measurement p_D	1.0	1.0	0.99	0.95
Filter p_D	0.99	0.99	0.99	0.95
Clustering U_c	36	100	100	100

Figure 4.5(a) provides the OSPA position errors for all cases. The top plot shows large OSPA errors for case ST1A throughout the observation window, as well as large initial errors

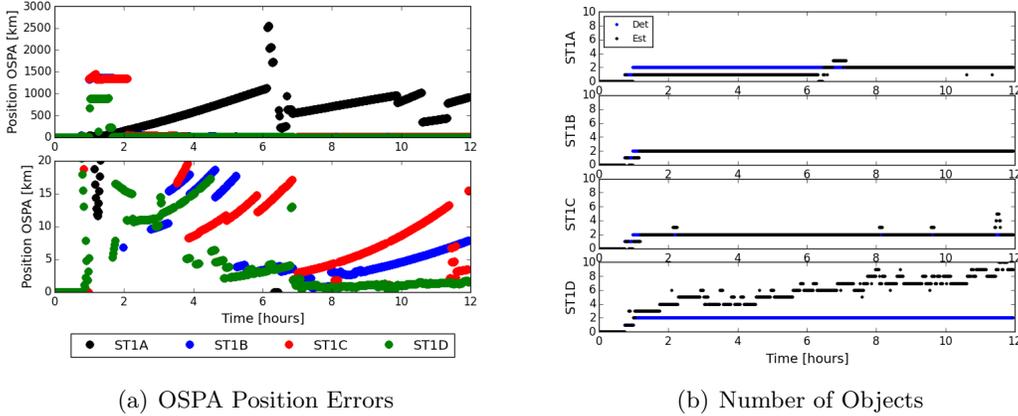


Figure 4.5: Single Thread CPHD Test Results

for the remaining three cases at the times the objects are first detected. The close up view in the lower plot demonstrates that for all cases with the larger clustering threshold $U_c = 100$, the filter converges to state estimate errors under 10 km by the final time. The increased clustering threshold also produces a better cardinality estimate, as shown in Figure 4.5(b). In the figure, the blue line indicates the number of objects that have been detected, while the black line represents the estimated cardinality. Case ST1B quickly converges on the correct number of targets and maintains its value as there are no missed detections or clutter returns. Case ST1C also performs well, despite the occurrence of five false alarms. Following each, the estimated cardinality increases as a result of the newly added CAR components, but these are quickly discarded and the filter returns to the correct value. Case ST1D overestimates the number of objects as a result of processing more than 40 false alarms during the course of the observation window.

Table 4.5 provides a summary of the test results, including the number of missed detections and false alarms. Note that the occurrence of two missed detections during the ST1D case causes no jump in the OSPA errors, an indication that the clustering technique is successfully mitigating the previously discussed spooky effect. All cases demonstrate the ability of the algorithm to locate and schedule follow-on measurements for objects with no *a priori* information, with the sensor spending the majority of its time in search mode.

Table 4.5: Test Case 1 Results

	ST1A	ST1B	ST1C	ST1D
Real Measurements	16	34	16	20
Missed Detections	0	0	0	2
False Alarms	0	0	5	41
Detected Objects	2	2	2	2
Estimated Cardinality	2	2	2	9
Search Mode (%)	98.3	95.83	97.64	93.61
Track Mode Real (%)	1.7	4.17	1.67	2.50
Track Mode Clutter (%)	0.0	0.0	0.69	3.89

Figure 4.6 gives a more detailed review of the ST1C test case. On the left, Figure 4.6(a) shows the number of measurements, missed detections, false alarms, and the sum of new target probabilities at each time. The sum of $p_{\text{new},i}$ values demonstrates the computed probability functions as desired, with values near 1 at the times corresponding to new targets and false alarms, and smaller values at times when measurements are collected near existing targets. On the right, Figure 4.6(b) provides insight on the tasking mode, including the expected information gain R_C , and the identity of the search grid point, object, or false alarm (denoted FA) scheduled for follow-on observation. The estimated number of objects is also included for reference.

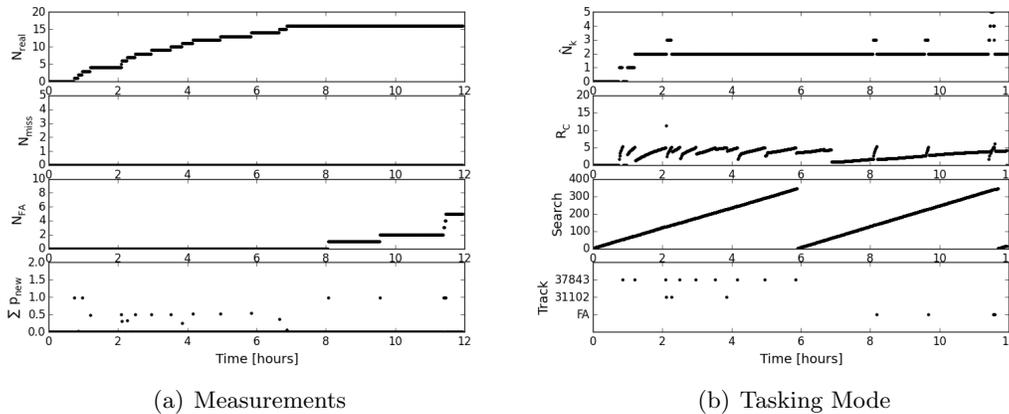


Figure 4.6: Test ST1C Results

While case ST1C is generally successful, increasing the number of false alarms causes issues in case ST1D. As shown in Figure 4.7, the filter overestimates the number of objects present due

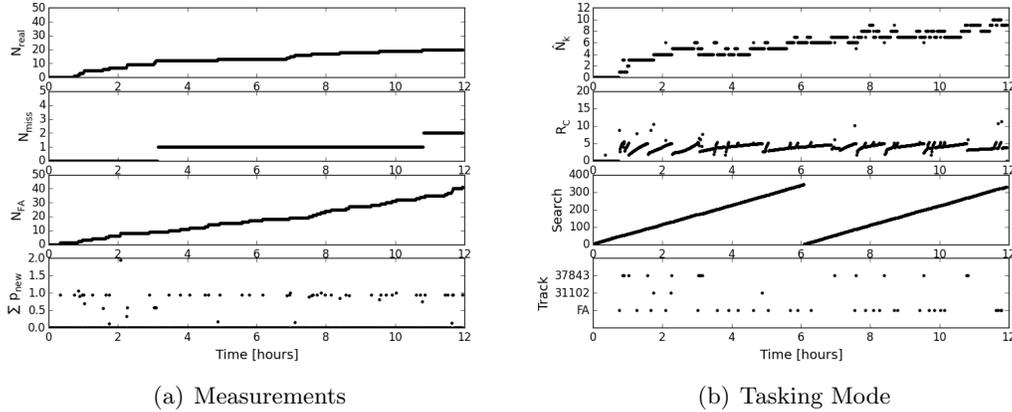


Figure 4.7: Test ST1D Results

to the large number of clutter returns in the measurement set. For each false alarm, a set of CAR GMM components with a collective weight close to one are added to the filter, and while the follow-on measurements prevent the estimated cardinality from growing too large, the filter is unable to recover the correct estimated target number as it did previously. In order to deal with increasingly challenging measurement conditions, subsequent test cases consider the use of a PHD filter to pre-screen measurements and refine CAR estimates prior to confirming targets for inclusion in the CPHD filter.

4.4.3 Test Case 2: Augmented CPHD with Missed Detections

The second test case examines the capabilities of the augmented CPHD filter, considering the effect of missed detections on the ability of the filter to estimate the multitarget state and cardinality. The same two objects are modeled as before, but over a time period of 36 hours to evaluate whether the algorithm can maintain custody of targets identified the first night following a 12-hour gap in sensor availability. Several values are considered for the probability of detection while no false alarms are simulated, again using $\lambda_{\kappa} = 0.01$ for the Poisson clutter model in both the PHD and CPHD filter updates. The information gain cutoff $C = 5$ and clustering threshold $U_c = 36$ are used for all cases. Clustering is applied in the confirmed CPHD filter only.

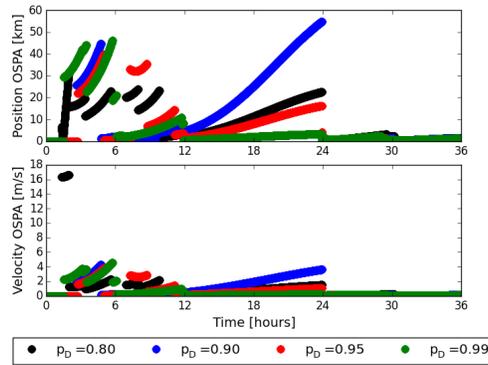


Figure 4.8: Missed Detection Test OSPA Results

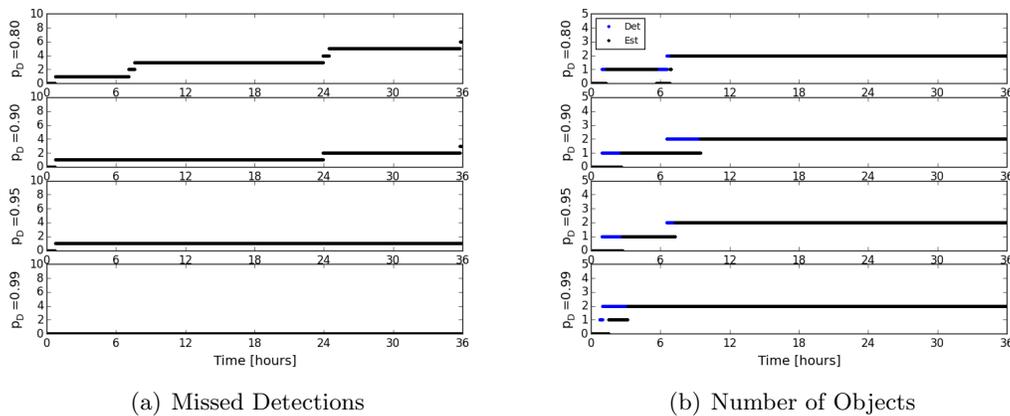


Figure 4.9: Missed Detection Test Results

Figures 4.8-4.9 provide the results of the missed detection test case. The OSPA position and velocity errors are given for the confirmed track CPHD filter, and for all cases converge to the km level by the final time. No unexpected peaks or jumps occur in the OSPA errors; the steady increase for most of the cases between 12 and 24 hours results from the sensor being inactive. Upon collecting new measurements at the 24-hour mark, the errors are quickly corrected.

Figure 4.9 provides information on the number of missed detections, as well as the number of detected and estimated objects, according to the MAP of the cardinality of the confirmed CPHD filter. The results demonstrate the reliable performance of the filter, correctly confirming exactly 2 objects in all cases, even in the presence of missed detections. Both the $p_D = 0.90$ and $p_D = 0.80$

cases experience missed detections at times when 2 objects are confirmed in the filter, but no large OSPA errors are observed as the spooky effect is mitigated by clustering.

In general, the filter performs well in terms of maintaining the estimated number of objects, though the $p_D = 0.80$ case drops one confirmed object around the 6-hour mark. The object is quickly recovered and reconfirmed. Since the occurrences of missed detections are highly variable and can produce different results, a total of 5 instances of each probability of detection case are simulated. The results presented here are for one instance in which measurements for all cases are generated by the same initial state for the random number generator, and are taken to be representative of the algorithm behavior in the presence of missed detections. The $p_D = 0.99$ case provides the expected results in an ideal scenario, in which no missed detections occur, which is true across all 5 instances of this case. Two of the $p_D = 0.80$ instances include an object dropped by the CPHD filter, though at the final time, both correctly estimate that two objects are present. No objects are dropped in any $p_D = 0.95$ or $p_D = 0.90$ case. Through all scenarios tested, knowledge of any confirmed objects are maintained across the 12 hour period of inactivity, allowing the system to schedule observations for known objects without having to rediscover them by scanning the search region. A summary of the SDT algorithm performance is provided in Table 4.6.

Table 4.6: Test Case 2 Results

	$p_D = 0.99$	$p_D = 0.95$	$p_D = 0.90$	$p_D = 0.80$
Real Measurements	26	25	24	24
Missed Detections	0	1	3	6
False Alarms	0	0	0	0
Detected Objects	2	2	2	2
Estimated Cardinality	2	2	2	2
Search Mode (%)	98.8	98.8	98.7	98.5
Track Mode Real (%)	1.2	1.2	1.3	1.5
Track Mode Clutter (%)	0.0	0.0	0.0	0.0

Figure 4.10 presents additional results from the $p_D = 0.90$ case. Figure 4.10(a) shows the number of measurements, missed detections, false alarms, and the sum of new target probabilities at each time. The sum of $p_{\text{new},i}$ values demonstrates the computed probability functions as desired,

with values near 1 at the times new objects are discovered, and generally close to 0 at times when measurements are collected near existing targets. The low values observed for measurements near existing targets are notably much better than for the single thread CPHD filter.

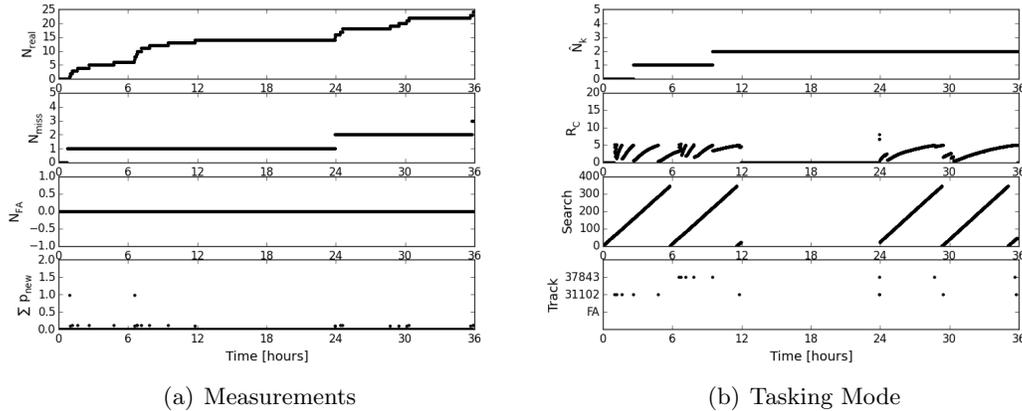


Figure 4.10: Missed Detection $p_D=0.90$ Test Results

Figure 4.10(b) provides details on the tasking mode, including the expected information gain, and the appropriate search bin or target ID scheduled depending on the selected sensor tasking mode. A missed detection occurs for object 37843 during the first search window, and therefore no follow-on tracking is scheduled until it is observed during the second pass through the search region. On the second night, with both objects already confirmed, the sensor continues to alternately observe each one to maintain the filter estimate.

4.4.4 Test Case 3: Augmented CPHD with False Alarms

The third test case considers the effect of false alarms on the ability of the filter to estimate the multitarget state and cardinality. The same two objects are modeled, again over a time period of 36 hours. Several values are considered for the Poisson clutter rate while no missed detections are simulated, using the value $p_D = 0.99$ for both the PHD and CPHD filter updates. The information gain cutoff $C = 5$ and clustering threshold $U_c = 36$ are used for all cases.

Figure 4.11 provides the OSPA results for the position and velocity estimates from the con-

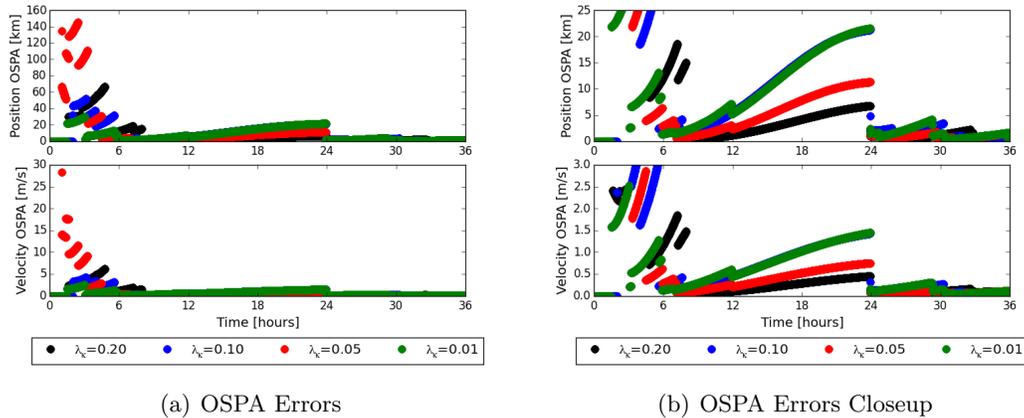


Figure 4.11: Clutter Test OSPA Results

firmed track CPHD filter, with the closeup view showing errors at the km level for all cases by the final time. No unexpected peaks or jumps occur in the OSPA errors, the steady increase for most cases between 12 and 24 hours coincides with the sensor being inactive. Figure 4.12 provides the number of false alarms and estimated number of objects for each case, from which it is clear the filter maintains the correct number of targets throughout, even for cases including hundreds of false measurements.

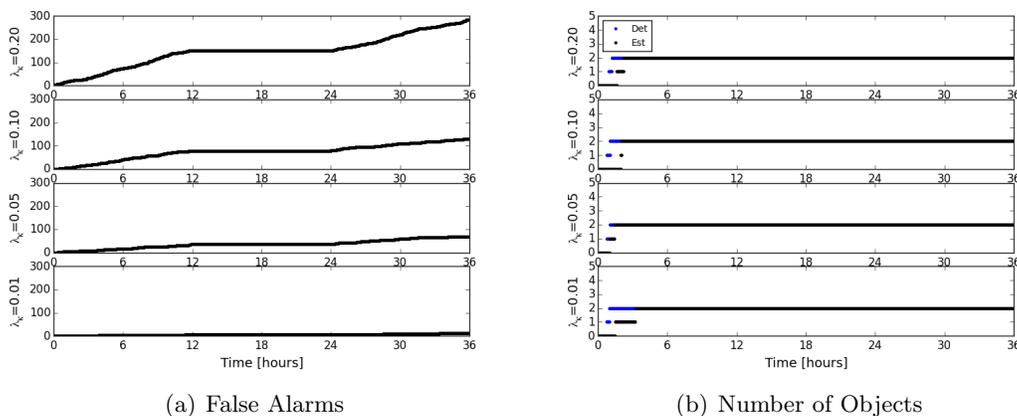


Figure 4.12: Clutter Test Results

The use of the PHD pre-filter is a substantial improvement in terms of mitigating the effects of clutter on the cardinality estimate, as expected because the confirmation process requires multiple

measurements for a given object before it is incorporated in the CPHD mixture. Assuming clutter is uniformly distributed in the field of view, it is unlikely for false alarms to occur in such a pattern as to produce a confirmed target. A summary of the test case results is provided in Table 4.7, showing an increasing amount of time spent in clutter tracking as the number of false alarms increases, though all cases spend the majority of time in search mode.

Table 4.7: Test Case 3 Results

	$\lambda_{\kappa} = 0.01$	$\lambda_{\kappa} = 0.05$	$\lambda_{\kappa} = 0.10$	$\lambda_{\kappa} = 0.20$
Real Measurements	23	23	22	26
Missed Detections	0	0	0	0
False Alarms	13	68	129	283
Detected Objects	2	2	2	2
Estimated Cardinality	2	2	2	2
Search Mode (%)	98.0	94.5	90.4	81.0
Track Mode Real (%)	1.0	1.0	1.0	1.2
Track Mode Clutter (%)	1.0	4.4	8.6	17.8

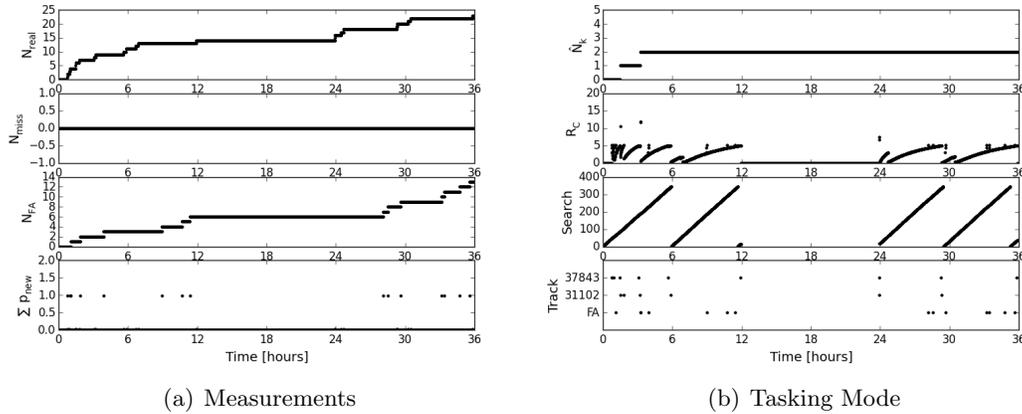
Figure 4.13: Clutter $\lambda_{\kappa} = 0.01$ Test Results

Figure 4.13 presents additional results from the $\lambda_{\kappa} = 0.01$ case. Figure 4.13(a) provides details on the measurements collected. The sum of $p_{\text{new},i}$ values demonstrate the computed probability functions as desired, with values near 1 corresponding to new objects or false alarms. In particular, at times when real measurements are collected, $p_{\text{new},i}$ is close to zero with the exception of the original first two detections. Figure 4.13(b) provides details on the expected information gain and

sensor tasking mode. Despite routinely following up on false alarms, the SDT algorithm still spends the majority of its time in search mode, while continuing to schedule observations of both confirmed objects to maintain well localized estimates of them following the gap in sensor availability.

4.4.5 Test Case 4: Augmented CPHD with Ten Objects

The final test case considers the effect of missed detections and false alarms while simulating all ten objects listed in Table 4.3. The value of the information gain cutoff is varied, while the probability of detection and clutter rates are held fixed at $p_D = 0.95$ and $\lambda_{\kappa} = 0.20$. The simulation is performed for a total of 60 hours, covering two 12-hour gaps in sensor availability.

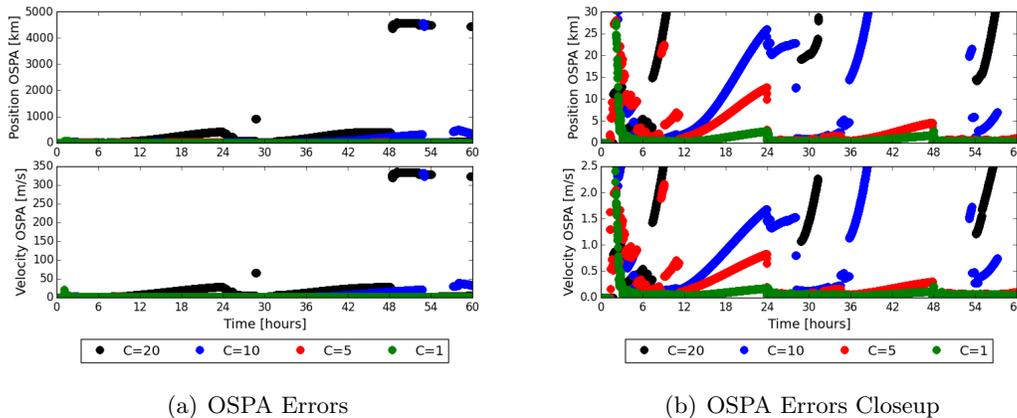


Figure 4.14: Ten Object OSPA Results

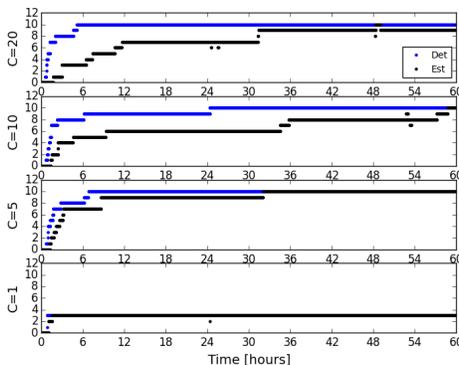
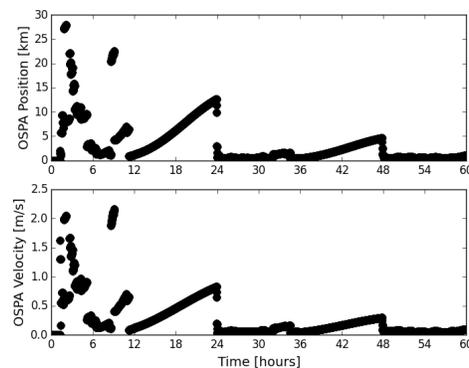


Figure 4.15: Number of Objects

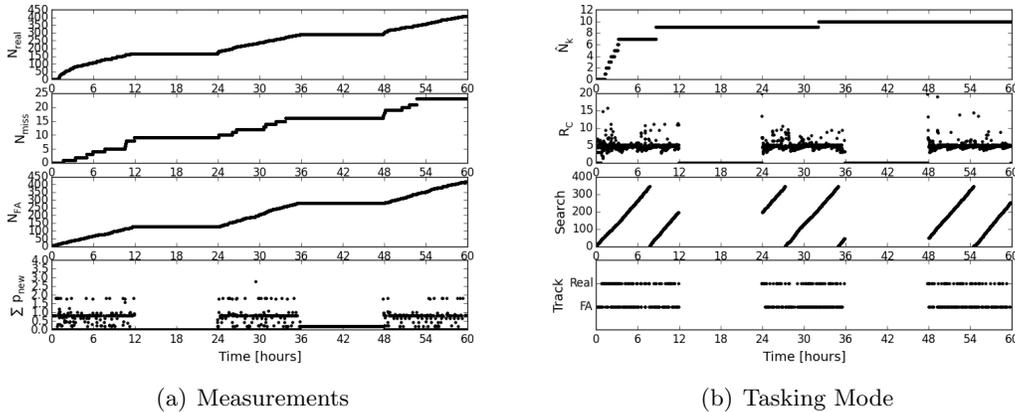
Figure 4.14 provides the OSPA results for the position and velocity estimates. From the closeup view, both the $C = 1$ and $C = 5$ cases perform well, achieving errors at the km level by the final time. The $C = 10$ and $C = 20$ cases produce large spikes in the OSPA errors during the final observation window, and in general, appear to be less reliable. Figure 4.15 provides the estimated number of objects for each case, and confirms the poor performance of these cases, which produce several occurrences of dropped objects. The $C = 1$ case gets stuck following up on the first few objects, and fails to complete the search grid even once. The $C = 5$ case performs well, however, as discussed in detail below. Results for all cases are summarized in Table 4.8.

Table 4.8: Test Case 4 Results

	$C = 1$	$C = 5$	$C = 10$	$C = 20$
Real Measurements	5952	406	172	94
Missed Detections	345	23	10	5
False Alarms	427	418	446	432
Detected Objects	3	10	10	10
Estimated Cardinality	3	10	10	9
Search Mode (%)	2.1	75.3	82.6	98.6
Track Mode Real (%)	96.9	8.3	2.6	0.6
Track Mode Clutter (%)	0.9	16.4	14.8	0.8

Figure 4.16: Ten Object $C = 5$ OSPA Errors

As the most successful case, the $C = 5$ results are examined more closely. The OSPA errors in Figure 4.16 demonstrate the desired behavior, eventually converging to the km level. Despite the fact that over 20 detections are missed and over 400 false alarms occur, the filter does not

Figure 4.17: Ten Object $C = 5$ Results

produce any unexpected peaks in OSPA errors due to shifting GMM weights, nor are any objects dropped or false alarm tracks incorrectly confirmed. The filter confirms 9 objects the first night, and maintains the correct estimate of 10 objects from the 32-hour mark forward. In total, the filter spends approximately 75% of its time in search mode, a good indication that the SDT algorithm could continue to find and maintain tracks of additional objects if needed. The time spent in tracking mode favors following up on clutter, though it is clear that follow-on measurements of real objects are scheduled regularly through all 3 observation windows, as shown in Figure 4.17(b).

4.4.6 Limitations of the Method

The results presented demonstrate a viable solution to the single-sensor SDT problem, in which a filter initialized with no *a priori* knowledge is able to find and track all objects in the scene in the presence of clutter and missed detections. The proposed solution does exhibit some limitations, and could be improved to approach the problem in a more general or abstract way. As a whole, the method is computationally complex, and involves a large parameter trade space. Fortunately, the parameters have different effects on the algorithm performance and can therefore be grouped and considered separately from one another, simplifying any tuning or analysis necessary. User-defined boundaries on the CAR and uncertainties used to generate the GMM approximation

from it are generally concerns to computational burden only. Thresholds for GMM component pruning, merging, and clustering affect the tendency of the filter to confirm or drop objects. Finally, the ad hoc information gain cutoff proposed to switch between search and track mode affects the amount of time spent in each and the ability of the algorithm to produce sufficient measurements to confirm all objects present.

A simplified analysis on the interpretation of information gain for the single object case is included in Appendix C.4, but it is too limited to provide a comprehensive approach for determining a useful switch point. A more principled approach would be to characterize the information gain associated with maintaining the multitarget errors of known objects to a certain precision, while also defining a discrete value based in information theory for instantiating a new track. The sensor allocation scheme could then weigh the relative benefit of catalog maintenance against the value of discovering new objects, assuming some probability of finding new targets in a given region. However, this still produces a solution that is highly dependent on the problem, as the number of objects in the field of view affects the expected information gain. A switch point defined to maintain a catalog of objects distributed through the search space such that only one is visible at a time may not be appropriate for a problem in which objects are grouped closely together. Generalizing the method to address these issues is considered a task for future study.

4.5 Chapter Summary

The theory and application of an IOD method based on the admissible region has been presented, as well as details of its use to instantiate new targets in the CPHD filter. The method was combined with previously discussed sensor allocation techniques to search for new objects and perform follow-on tracking to improve and maintain estimates of targets after their inclusion in the filter. Simulation results demonstrate the ability of the method to find and track objects with no *a priori* information, other than the assumption that they are in geostationary orbits visible to the sensor. Despite the limitations discussed, the method generally localizes targets well enough to maintain custody of them following gaps in sensor availability, and use of the PHD

augmented filter provides a degree of robustness estimating the target number in the presence of sparse clutter and missed detections. The combination of the IOD, filtering, and sensor allocation methods demonstrates the ability to solve the search-detect-track problem within a FISST-derived framework.

Chapter 5

Conclusion

5.1 Research Summary

This dissertation presented results from the application of Finite Set Statistics filtering methods to the problem of estimating geosynchronous space objects. Chapter 2 provided background and motivation for the problem and examined the use of several multitarget filters for GEO SSA. In particular, the CPHD filter was determined to be a suitable approach, a conclusion further exemplified by the study in Chapter 3, in which a catalog of nearly 1000 objects was maintained using simulated measurements from three ground stations. The proposed sensor allocation scheme based on maximizing the Rényi divergence produced superior performance as compared to simpler grid-based search methods, and both single-step and multistep approaches were deemed successful. Single-step sensor tasking offers the additional advantage of being able to quickly retask sensors, a feature utilized in Chapter 4 to perform follow-on tracking for newly discovered objects.

The primary contribution of this research is the application of the information gain reward function for sensor allocation. Through simple demonstration, the choice of Rényi divergence was shown to be a more appropriate objective function than the Cauchy-Schwarz divergence for tasking in which a subset of objects are scheduled for observation. This conclusion confirmed previous research studying the use of relative and absolute information gain functionals to drive sensor tasking for SSA. The large scale simulation included a realistic set of objects from the TLE catalog, on the same order of magnitude as the number of objects currently tracked at GEO. The results indicate the benefits of efficient sensor tasking, which allows for better catalog maintenance

and alleviates the sparse data condition that drives many other challenges in SSA. Additional constraints on real-world sensors driven by local weather conditions or the apparent magnitudes of space objects further emphasize the need for efficient measurement collection.

Use of information theoretic tasking was explored further in Chapter 4, which used the Rényi divergence to drive follow-on tracking for newly discovered objects. Implementation of a measurement-based birth model for the CPHD filter was discussed using the CAR IOD technique. The standard CPHD filter was augmented by a PHD mixture to maintain unconfirmed tracks, and a method was developed to approximate the probability that measurements represented new targets. The combination of these techniques was successful in finding and tracking a small number of geostationary objects in the presence of sparse clutter and missed detections.

5.2 Future Research Considerations

The findings presented in this research suggest a number of possible improvements and directions for future research. As discussed in Chapter 4, one of the drawbacks of the CPHD filter is its tendency to shift probability mass from undetected to detected objects, also known as the spooky effect or spooky action at a distance. Augmenting the CPHD filter with a clustering technique alleviated this issue [59], but the behavior suggests better alternatives may be available. Namely, the multi-Bernoulli filter is not impacted by this issue. The latest incarnation, the δ -Generalized Labeled Multi-Bernoulli (δ -GLMB) filter, resembles track-oriented MHT, and may provide a better approach to SSA than the CPHD filter while still utilizing the full rigor of Finite Set Statistics as a mathematical framework [81, 82]. In general, the field of multitarget estimation is still evolving rapidly, and new techniques may provide advantages in representing the orbital debris catalog and filtering noisy and ambiguous measurements.

Many simplifying assumptions were employed in the sensor allocation study, most fundamentally the implicit assumption that the sensors were centrally networked and operating from a shared object catalog, which was updated in real time. New research has examined the use of distributed control sensor networks and information fusion using FISST [78], which would be highly applicable

to the SSA problem. The use of the Rényi divergence as an objective function is computationally burdensome due to the need for a numeric solution of the integral. The Cauchy-Schwarz divergence would make a better option if it can be modified to properly compare the expected information gain for objects with different initial covariances. As a measure of absolute information gain, it may also lead to a more principled technique to optimize sensor tasking between searching for new objects and maintaining the catalog of known objects. To initiate tracks for new objects, the CAR IOD technique is suitable for the 4D measurement set presented, but may not extend easily to other measurement types, such as range and range-rate. A more general solution to the IOD problem for multitarget filtering, for instance using a solution to the multiscan assignment problem [60], may be more useful.

Finally, the incorporation of target priorities in the sensor tasking objective function is an important consideration, both to reflect the needs of the SSA user community and to solve pressing issues such as collision avoidance. In particular, a predicted collision should cause the system to increase the tasking priorities of the objects involved, prompting additional measurements and refinement of the prediction. The goal in sensor allocation should be to automate the process as much as possible, as the increasing size of the catalog mandates efficient collection of measurements.

Bibliography

- [1] Jeffrey M. Aristoff, Joshua T. Horwood, Navraj Singh, Aubrey B. Poore, Carolyn Sheaff, and M.K. Jah. Multiple hypothesis tracking (mht) for space surveillance: Theoretical framework. In AAS/AIAA Astrodynamics Specialist Conference, Hilton Head, SC, August 2013.
- [2] Michael Beard, Ba-Tuong Vo, Ba-Ngu Vo, and Sanjeev Arulampalam. A partially uniform target birth model for gaussian mixture phd/cphd filtering. IEEE Transactions on Aerospace and Electronic Systems, 49(4):2835–2844, October 2013.
- [3] Michael Beard, Ba-Tuong Vo, Ba-Ngu Vo, and Sanjeev Arulampalam. Sensor control for multi-target tracking using cauchy-schwarz divergence. In Information Fusion (FUSION), 2015 18th International Conference on, pages 937–944, Washington, D.C., July 2015.
- [4] Samuel Blackman. Multiple Target Tracking with Radar Applications. Artech House, Norwood, MA, 1986.
- [5] Samuel Blackman and R. Popoli. Design and Analysis of Modern Tracking Systems. Artech House, Norwood, MA, 1999.
- [6] Samuel S. Blackman. Multiple hypothesis tracking for multiple target tracking. IEEE Aerospace and Electronic Systems Magazine, Part 2: Tutorials, 19(1):1–18, January 2004.
- [7] Ben Bradley and Penina Axelrad. Improved estimation of orbits and physical properties of objects in geo. In Advanced Maui Optical and Space Surveillance Technologies Conference (AMOS), Maui, HI, September 2013.
- [8] Daniel S. Bryant, Emmanuel D. Delande, Steven Gehly, Jeremie Houssineau, Daniel E. Clark, and Brandon A. Jones. The cphd filter with target spawning. IEEE Transactions on Signal Processing, 2015. T-SP-19756-2015.R3.
- [9] S. Challa, B. Vo, and X. Wang. Bayesian approaches to track existence - ipda and random sets. In Information Fusion (FUSION), 2002 5th International Conference on, Annapolis, MD, July 2002.
- [10] Chia-Chun Chao. Applied Orbit Perturbation and Maintenance. The Aerospace Press, El Segundo, CA, 2005.
- [11] Y. Cheng, K.J. DeMars, C. Fruh, and M.K. Jah. Gaussian mixture phd filter for space object tracking. In AAS/AIAA Space Flight Mechanics Meeting, Kauai, Hawaii, February 2013.

- [12] D. E. Clark, J. Houssineau, and E. Delande. A Few Calculus Rules for Chain Differentials. 2015. arXiv:1506.08626.
- [13] Daniel E. Clark and Jeremie Houssineau. Faá di bruno’s formula and spatial cluster modeling. *Spacial Statistics*, 6:109–117, 2013.
- [14] Ingemar J. Cox and Sunita L. Hingorani. An efficient implementation of reid’s multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(2):136–150, February 1996.
- [15] Emmanuel Delande, Murat Üney, Jérémie Houssineau, and Daniel E. Clark. Regional variance for mult-object filtering. *IEEE Transactions on Signal Processing*, 62(13):3415–3428, July 2014.
- [16] K. J. DeMars, I. I. Hussein, M. K. Jah, and R. S. Erwin. The cauchy-schwarz divergence for assessing situational information gain. In *Information Fusion (FUSION), 2012 15th International Conference on*, pages 1126–1133, July 2012.
- [17] Kyle J. DeMars. *Nonlinear Orbit Uncertainty Prediction and Rectification for Space Situational Awareness*. PhD thesis, University of Texas at Austin, Austin, TX, 2010.
- [18] 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, July-August 2013.
- [19] 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, September-October 2013.
- [20] P. R. Escobal. *Methods of Orbit Determination*. Krieger, Malabar, FL, 1965.
- [21] Walter J. Faccenda, David Ferris, C. Max Williams, and Dave Brisnehan. Deep stare technical advancements and status. In *Advanced Maui Optical and Space Surveillance Technologies Conference*, Maui, HI, September 2003.
- [22] Tim Flohrer and S. Frey. Classification of geosynchronous objects. *Issue 18 ESA/ESOC*, 2016.
- [23] William M. Folkner, James G. Williams, Dale H. Boggs, Ryan S. Park, and Petr Kuchynka. The planetary and lunar ephemerides de430 de431. IPN Progress Report 42-196, Jet Propulsion Laboratory, California Institute of Technology, http://ipnpr.jpl.nasa.gov/progress_report/42-196/196C.pdf, February 2009.
- [24] Thomas E. Fortmann, Yaakov Bar-Shalom, and Molly Scheffe. Sonar tracking of multiple targets using joint probabilistic data association. *IEEE Journal of Oceanic Engineering*, OE-8(3):173–184, July 1983.
- [25] D. Franken and M. Ulmke. ‘spooky action at a distance’ in the cardinalized probability hypothesis density filter. *IEEE Transactions on Aerospace and Electronic Systems*, 45(4):1657–1664, 2009.
- [26] Kohei Fujimoto and Daniel J. Scheeres. Correlation of optical observations of earth-orbiting objects and initial orbit determination. *Journal of Guidance, Control, and Dynamics*, 35(1):208–221, 2012.

- [27] Steven Gehly, Brandon Jones, and Penina Axelrad. Comparison of multitarget filtering methods as applied to space situational awareness. In AAS/AIAA Astrodynamics Specialist Conference, Hilton Head, SC, August 2013.
- [28] Arthur Gelb. Applied Optimal Estimation. M.I.T. Press, Cambridge, MA, 1974.
- [29] R. H. Gooding. A new procedure for the solution of the classical problem of minimal orbit determination from three lines of sight. Celestial Mechanics and Dynamical Astronomy, 66:387–423, 1997.
- [30] Alfred O. Hero, Chris Kreucher, and Doron Blatt. Information theoretic approaches to sensor management. In A. O. Hero, D. A. Castanón, D. Cochran, and K. Kastella, editors, Foundations and Applications of Sensor Management, pages 33–57. Springer, New York, 2008.
- [31] Y. C. Ho and R. Lee. A bayesian approach to problems in stochastic estimation and control. IEEE Transactions on Automatic Control, 9(4):333–339, October 1964.
- [32] H. G. Hoang, B.-N. Vo, B.-T. Vo, and R. Mahler. The cauchy-schwartz divergence for poisson point processes. IEEE Transactions on Information Theory, 61(8):4475–4485, 2015.
- [33] H. G. Hoang and B.-T. Vo. Sensor management for multi-target tracking via multi-bernoulli filtering. Automatica, 50(4):1135–1142, 2014.
- [34] I.I. Hussein, K.J. DeMars, C. Fruh, R.S. Erwin, and M.K. Jah. An aegis-fisst integrated detection and tracking approach to space situational awareness. In Information Fusion (FUSION), 2012 15th International Conference on, pages 2065–2072, July 2012.
- [35] Brandon A. Jones, Steven Gehly, and Penina Axelrad. Measurement-based birth model for a space object cardinalized probability hypothesis density filter. In AIAA/AAS Astrodynamics Specialist Conference, San Diego, CA, August 2014.
- [36] Simon J. Julier and Jeffrey K. Uhlmann. New extension of the kalman filter to nonlinear systems. In Proc. SPIE 3068, Signal Processing, Sensor Fusion, and Target Recognition VI, pages 182–193, Orlando, 1997.
- [37] Simon J. Julier and Jeffrey K. Uhlmann. Unscented filtering and nonlinear estimation. Proceedings of the IEEE, 92(3):401–422, March 2004.
- [38] T. Kirubarajan, H. Wang, Y. Bar-Shalom, and K. R. Pattipatti. Efficient multisensor fusion using multidimensional data association. IEEE Transactions on Aerospace Systems, 37(2):386–398, April 2001.
- [39] H. Krag, M. Kahl, J. Bendisch, H. Klinkrad, and T. Schildknecht. Space based optical observation of small debris objects. In Proceedings of the Third European Conference on Space Debris, volume 1 of ESA SP-473, pages 147–152, Darmstadt, Germany, March 19-21 2001.
- [40] S. Kullback and R. A. Leibler. On information and sufficiency. The Annals of Mathematical Statistics, 22(1):79–86, March 1951.
- [41] T. Kurien. Issues in the design of practical multitarget tracking algorithms. In Y. Bar-Shalom, editor, Multitarget-Multisensor Tracking: Advanced Applications, chapter 3. Artech House, Norwood, MA, 1990.

- [42] Chee Sing Lee, Daniel E. Clark, and Joaquim Salvi. Slam with dynamic targets via single-cluster phd filtering. IEEE Journal of Selected Topics in Signal Processing, 7(3):543–552, June 2013.
- [43] L. Lin, Y. Bar-Shalom, and T. Kirubarajan. Track labeling and phd filter for multitarget tracking. IEEE Transactions on Aerospace and Electronic Systems, 42(3):778–795, July 2006.
- [44] Ronald Mahler. Multitarget sensor management of dispersed mobile sensors. In D. Grundel, R. Murphey, and P. Pardalos, editors, Theory and Algorithms for Cooperative Systems, pages 239–310. World Scientific Publishing Co., Singapore, 2004.
- [45] Ronald P. S. Mahler. Multitarget bayes filtering via first-order multitarget moments. IEEE Transactions on Aerospace and Electronic Systems, 39(4):1152–1178, October 2003.
- [46] Ronald P. S. Mahler. Sensor management with non-ideal sensor dynamics. In Information Fusion (FUSION), 2004 7th International Conference on, Stockholm, Sweden, June-July 2004.
- [47] Ronald P. S. Mahler. 'statistics 101' for multisensor, multitarget data fusion. IEEE AE Systems Magazine, 1(1):53–64, January 2004.
- [48] Ronald P. S. Mahler. Phd filters of higher order in target number. IEEE Transactions on Aerospace and Electronic Systems, 43(4):1523–1543, October 2007.
- [49] Ronald P. S. Mahler. Statistical Multisource-Multitarget Information Fusion. Artech House, Norwood, MA, 2007.
- [50] Ronald P. S. Mahler. Advances in Statistical Multisource-Multitarget Information Fusion. Artech House, Norwood, MA, 2014.
- [51] Jared M. Maruskin, Daniel J. Scheeres, and Kyle T. Alfriend. Correlation of optical observations of objects in earth orbit. Journal of Guidance, Control, and Dynamics, 32(1):194–209, January-February 2009.
- [52] A. Milani, G. F. Gronchi, Mattia De' Michieli Vitturi, and Zoran Knežević. Orbit determination with very short arcs. i admissible regions. Celestial Mechanics and Dynamical Astronomy, 90(1-2):57–85, September 2004.
- [53] K. G. Murty. An algorithm for ranking all the assignments in order of increasing cost. Operations Research, 16:682–687, 1968.
- [54] NASA. Orbital debris quarterly news volume 20, issue 3. Technical report, NASA Orbital Debris Program Office, July 2016.
- [55] M. Oswald, H. Krag, P. Wegener, and B. Bischof. Concepts for an orbital telescope observing the debris environment in geo. Advances in Space Research, 34:1155–1159, 2004.
- [56] K. Panta, B.-N. Vo, and S. Singh. Novel data association schemes for the probability hypothesis density filter. IEEE Transactions on Aerospace and Electronic Systems, 43(2):556–570, April 2007.
- [57] Kusha Panta, Daniel E. Clark, and Ba-Ngu Vo. Data association and track management for the gaussian mixture probability hypothesis density filter. IEEE Transactions on Aerospace and Electronic Systems, 45(3):1003–1016, July 2009.

- [58] Nikolaos K. Pavlis, Simon A. Holmes, Steve C. Kenyon, and John K. Factor. An Earth gravitational model to degree 2160: EGM2008. In Proceedings of the European Geosciences Union General Assembly, Vienna, Austria, April 13-18 2008.
- [59] Y. Petetin, D. Clark, B. Ristic, and D. Maltese. A tracker based on a cphd filter approach for infrared applications. In Proc. of SPIE, volume 8050, 2011.
- [60] Aubrey B. Poore. Multidimensional assignment formulation of data association problems arising from multitarget and multisensor tracking. Computational Optimization and Applications, 3:27–57, 1994.
- [61] Aubrey B. Poore and Alexander J. Robertson III. A new lagrangian relaxation based algorithm for a class of multidimensional assignment problems. Computational Optimization and Applications, 8:129–150, 1996.
- [62] G. W. Pulford. Taxonomy of multiple target tracking methods. IEE Proceedings on Radar and Sonar Navigation, 152(5):291–304, October 2005.
- [63] Donald B. Reid. An algorithm for tracking multiple targets. IEEE Transactions on Automatic Control, 24(6):843–854, December 1979.
- [64] Alfréd Rényi. On measures of entropy and information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics, pages 547–561, Berkeley, CA, 1961. University of California Press.
- [65] B. Ristic, D. Clark, B.-N. Vo, and B.-T. Vo. Adaptive target birth intensity for phd and cphd filters. IEEE Transactions on Aerospace and Electronic Systems, 48(2):1656–1668, 2012.
- [66] B. Ristic, B.-N. Vo, and D. Clark. A note on the reward function for phd filters with sensor control. IEEE Transactions on Aerospace and Electronic Systems, 47(2):1521–1529, 2011.
- [67] T. Schildknecht, T. Flohrer, R. Musci, and R. Jehn. Statistical analysis of the esa optical space debris surveys. Acta Astronautica, 63:119–127, 2008.
- [68] Dominic Schuhmacher, Ba-Tuong Vo, and Ba-Ngu Vo. A consistent metric for performance evaluation of mult-object filters. IEEE Transactions on Signal Processing, 56(8):3447–3457, August 2008.
- [69] Robert Scott and Brad Wallace. Small aperture telescope observations of co-located geostationary satellites. In Advanced Maui Optical and Space Surveillance Technologies Conference, Maui, HI, September 2009.
- [70] Claude E. Shannon. A mathematical theory of communication. Bell System Technical Journal, 27(3):379–423, 1948.
- [71] E. Soop. Handbook of Geostationary Orbits. Kluwer Academic, Dordrecht, 1994.
- [72] H. W. Sorenson and D. L. Alspach. Recursive bayesian estimation using gaussian sums. Automatica, 7:465–479, 1971.
- [73] J. Stauch, M. Jah, J. Baldwin, T. Kelecy, and K. A. Hill. Mutual application of joint probabilistic data association, filtering, and smoothing techniques for robust multiple space object tracking (invited). In AIAA/AAS Astrodynamics Specialist Conference, San Diego, CA, 2014.

- [74] Byron Tapley, Bob Schutz, and George Born. Statistical Orbit Determination. Elsevier Academic Press, New York, first edition, 2004.
- [75] Jill Tombasco. Orbit Estimation of Geosynchronous Objects via Ground-based and Space-based Optical Tracking. PhD thesis, University of Colorado at Boulder, Boulder, CO, 2011.
- [76] G. Tommei, A. Milani, and A. Rossi. Orbit determination of space debris: admissible regions. Celestial Mechanics and Dynamical Astronomy, 97:289–304, 2007.
- [77] M. Ulmke, Ozgur Erdinc, and Peter Willett. Gaussian mixture cardinalized phd filter for ground moving targets. In Proceedings 10th International Conference on Information Fusion, Quebec, Canada, July 2007.
- [78] Murat Uney, Daniel E. Clark, and Simon J. Julier. Distributed fusion of phd filters via exponential mixture densities. IEEE Journal of Selected Topics in Signal Processing, 7(3):521–531, 2013.
- [79] David A. Vallado. Fundamentals of Astrodynamics and Applications. Microcosm Press, Hawthorne, CA, third edition, 2007.
- [80] Ba-Ngu Vo and Wing-Kin Ma. The gaussian mixture probability hypothesis density filter. IEEE Transactions on Signal Processing, 54(11):4091–4104, 2006.
- [81] Ba-Ngu Vo, Ba-Tuong Vo, and D. Phung. Labeled random finite sets and the bayes multi-target tracking filter. IEEE Transactions on Signal Processing, 62(24):6554–6567, 2014.
- [82] Ba-Tuong Vo and Ba-Ngu Vo. Labeled random finite sets and multi-object conjugate priors. IEEE Transactions on Signal Processing, 61(13):3460–3475, July 2013.
- [83] Ba-Tuong Vo, Ba-Ngu Vo, and Antonio Cantoni. Analytic implementations of the cardinalized probability hypothesis density filter. IEEE Transactions on Signal Processing, 55(7):3553–3567, 2007.
- [84] Ba-Tuong Vo, Ba-Ngu Vo, and Antonio Cantoni. The cardinality balanced multi-target multi-bernoulli filter and its implementations. IEEE Transactions on Signal Processing, 57(2):409–423, February 2009.
- [85] Eric A. Wan and Rudolph Van Der Merwe. The unscented kalman filter for nonlinear estimation. In Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000, pages 153–158, Lake Louise, Alberta, October 2000.
- [86] Patrick S. Williams. Coupling Between Nonlinear Estimation and Dynamic Sensor Tasking Applied to Satellite Tracking. PhD thesis, Pennsylvania State University, State College, PA, 2012.
- [87] Patrick S. Williams, David B. Spencer, and Richard S. Erwin. Comparison of two single-step, myopic sensor management decision processes applied to space situational awareness. In Proceedings of the 22nd AAS/AIAA Spaceflight Mechanics Meeting, pages 173–190, Charleston, SC, January-February 2012.

- [88] Kevin Wyffels and Mark Campbell. Negative information for occlusion reasoning in dynamic extended multiobject tracking. IEEE Transactions on Robotics, 31(2):425–442, April 2015.
- [89] Tim Zajic and Ronald P. S. Mahler. Practical information-based data fusion performance evaluation. In Signal Processing, Sensor Fusion, and Target Recognition VIII, SPIE Proc., volume 3720, pages 92–103, Orlando, FL, July 1999.

Appendix A

Nomenclature

This appendix provides a guide to some of the notation used throughout this dissertation. Some symbols, particularly greek letters, are reused in different chapters and sections, though the quantities represented should be clear from context.

A.1 Single Target Estimation

$A(\mathbf{x}, t)$	=	linearized dynamics matrix
$\tilde{\alpha}, \tilde{m}, \tilde{\sigma}$	=	AEGIS splitting library weight, mean, standard deviation
$F(\mathbf{x}, \mathbf{u}, t), f(\mathbf{x})$	=	single target nonlinear dynamics function, Markov transition density
$G(\mathbf{x}, t), g(\mathbf{x})$	=	single target nonlinear measurement likelihood function, density
$H(\mathbf{x})$	=	differential entropy
J, w	=	number of GMM components, component weight
K, P_{xz}, P_{zz}	=	Kalman gain, cross-correlation covariance, innovation covariance
$\Lambda, V, \lambda_k, \mathbf{v}_k$	=	eigenvalue matrix, eigenvector matrix, k-th eigenvalue, eigenvector
$L, W, \alpha, \beta, \kappa, \lambda, \gamma$	=	state dimension, UKF weights, parameters
$p(\mathbf{x}), p_g(\mathbf{x}; \mathbf{m}, P)$	=	single target PDF, Gaussian PDF, mean \mathbf{m} , covariance P
$\bar{P}_k, P_{k k-1}$	=	<i>a priori</i> estimated state covariance
P_k	=	<i>a posteriori</i> estimated state covariance
σ	=	standard deviation
\mathbf{x}, \mathbf{z}	=	single target state, measurement

ϵ, R	=	measurement noise, measurement noise covariance
$\bar{\mathbf{x}}_k, \mathbf{x}_{k k-1}$	=	<i>a priori</i> estimated state vector
$\hat{\mathbf{x}}_k, \mathbf{x}_k$	=	<i>a posteriori</i> estimated state vector
χ, ζ	=	state sigma point matrix, measurement sigma point matrix
\mathbf{v}, Q	=	zero mean white noise process, process noise covariance
\bar{z}	=	predicted measurement
$z_{1:k}$	=	time series of measurements z_1, \dots, z_k

A.2 Multitarget Estimation

d_{ij}	=	association distance
$G[h]$	=	multitarget probability generating functional (PGFL)
J, w	=	number of GMM components, component weight
λ, λ_κ	=	clutter number density, Poisson mean number of clutter returns
N_k, N_{FOV}	=	estimated number of targets, number of targets in field of view
$p(X), \beta_X(S)$	=	multitarget PDF, belief-mass function
$p(n), p_\Gamma(n), p_\kappa(m)$	=	cardinality distribution, birth cardinality, clutter cardinality
p_S, p_D	=	probability of survival, probability of detection
$s(\mathbf{x}), j^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\})$	=	single-object spatial density, n-th order Janossy density
$\sigma_j(\cdot)$	=	elementary symmetric function
t	=	time
\mathbf{u}	=	process noise vector
$\nu(\mathbf{x}), \gamma(\mathbf{x}), \kappa(\mathbf{z})$	=	PHD function, birth PHD, clutter PHD
X, Z	=	multitarget state, multitarget measurement set
U, T	=	merging threshold, pruning threshold
V_s	=	volume of sensor FOV
$\mathbf{y}_k^{(j)}$	=	weighted innovation

A.3 Orbits and Measurements

$a, e, i, \Omega, \omega, M, n_m$	=	Keplerian orbital elements, mean motion
h, a_e, e_e, p_e, N_e	=	geodetic height, Earth reference ellipsoid parameters
\mathcal{E}, \mathbf{h}	=	specific orbital energy, angular momentum
\mathbf{q}	=	ground station position vector
\mathbf{r}, r	=	position vector, orbit radius
μ	=	gravitational parameter
x, y, z	=	position coordinates
α, δ	=	topocentric right ascension, declination
λ, ϕ	=	geodetic longitude, latitude
Λ, Φ	=	ENU spherical coordinates
$\hat{\rho}, \rho$	=	LOS unit vector, range

A.4 Sensor Allocation

α	=	Rényi divergence parameter
$C(\cdot), R(\cdot)$	=	Cauchy-Schwarz divergence, Rényi divergence
$f_0(X), f_1(X)$	=	prior and posterior multitarget PDFs
$p_{\text{new}}(\mathbf{z}), p_{\text{exist}}(\mathbf{z}), p_{\text{FA}}(\mathbf{z})$	=	probability of new target, existing target, false alarm
\mathbf{u}	=	sensor control vector

A.5 Index Conventions

i, j, k	=	measurement, GMM component, time index
m, n	=	measurement index, target number index
l, q	=	integer indices

Appendix B

Filter Algorithms

This appendix details the implementation of the AEGIS PHD and CPHD filters. The steps given here comprise the basic forms of the filters. They assume constant values of p_D and p_S that are the same for all objects, do not model target birth, and do not perform clustering to mitigate effects of shifting GMM weights due to missed detections.

Equations of Motion:

$$\dot{\mathbf{x}} = F(\mathbf{x}, \mathbf{u}, t) \tag{B.1}$$

Linear time rate of change of differential entropy:

$$\dot{H}(\mathbf{x}) = \text{trace}\{A(\mathbf{x}(t), t)\} \tag{B.2}$$

$$A(\mathbf{x}(t), t) = \left[\frac{\partial F}{\partial \mathbf{x}} \right]_{\mathbf{x}=\hat{\mathbf{x}}} \tag{B.3}$$

Measurement model:

$$\mathbf{z}_k = G(\mathbf{x}_k, t_k) + \boldsymbol{\epsilon}_k \tag{B.4}$$

B.1 AEGIS PHD Filter

1. Initial GMM approximation of PHD at time t_0 with track labels T_0 .

$$\nu_0(\mathbf{x}_0) \approx \sum_{j=1}^{J_0} w_0^{(j)} p_g(\mathbf{x}_0; \hat{\mathbf{x}}_0^{(j)}, P_0^{(j)})$$

$$T_0 = \{\tau_0^{(j)}\}_{j=1}^{J_0}$$

2. Compute sigma points and entropy for each component.

$$\chi_0^{(j)} = \begin{bmatrix} \hat{\mathbf{x}}_0^{(j)} & \hat{\mathbf{x}}_0^{(j)} + \gamma\sqrt{P_0^{(j)}} & \hat{\mathbf{x}}_0^{(j)} - \gamma\sqrt{P_0^{(j)}} \end{bmatrix}_{L \times (2L+1)} \quad (\text{B.5})$$

$$H_0^{(j)} = \frac{1}{2} \log |2\pi e P_0^{(j)}| \quad (\text{B.6})$$

3. Prediction Step $t \in [t_{k-1}, t_k]$

3.A. Step through time in increments of $\Delta t_s \leq \Delta t_k$. For $t \in [t_{s-1}, t_s]$, propagate the sigma points and linear PDF entropy for the j -th component to compute $\bar{\chi}_s^{(j)}$ and $\bar{H}_{\text{lin},s}^{(j)}$.

$$\dot{\chi}_{s-1}^{(j)} = F(\chi_{s-1}^{(j)}, t_{s-1}), \quad \chi^{(j)}(t_{s-1}) = \chi_{s-1}^{(j)} \quad (\text{B.7})$$

$$\dot{H}_{\text{lin}}^{(j)} = \text{trace}\{A_{s-1}^{(j)}\}, \quad H_{\text{lin}}^{(j)}(t_{s-1}) = H_{\text{lin},s-1}^{(j)} \quad (\text{B.8})$$

3.B. Compute the component *a priori* mean and covariance, retain track label.

$$\bar{\mathbf{x}}_s^{(j)} = \sum_{l=0}^{2L} W_l^m \bar{\chi}_{s,l}^{(j)} \quad (\text{B.9})$$

$$\bar{P}_s^{(j)} = \sum_{l=0}^{2L} W_l^c (\bar{\chi}_{s,l}^{(j)} - \bar{\mathbf{x}}_s^{(j)}) (\bar{\chi}_{s,l}^{(j)} - \bar{\mathbf{x}}_s^{(j)})^T \quad (\text{B.10})$$

$$\bar{\tau}_s^{(j)} = \tau_{s-1}^{(j)} \quad (\text{B.11})$$

3.C. Compute the entropy from the nonlinear PDF propagation.

$$\bar{H}_{\text{nl},s}^{(j)} = \frac{1}{2} \log |2\pi e \bar{P}_s^{(j)}| \quad (\text{B.12})$$

3.D. Check for nonlinearity. If $|\bar{H}_{\text{nl},s}^{(j)} - \bar{H}_{\text{lin},s}^{(j)}| > T_{\text{split}}$, split along the r -th eigenvector (corresponding to the largest eigenvalue).

3.D.i. Replace the j -th component with Q new components, retain track label of j -th component.

$$\bar{w}_s^{(j)} p_g(\mathbf{x}_s; \bar{\mathbf{x}}_s^{(j)}, \bar{P}_s^{(j)}) \approx \sum_{q=1}^Q \bar{w}_s^{(q)} p_g(\mathbf{x}_s; \bar{\mathbf{x}}_s^{(q)}, \bar{P}_s^{(q)}) \quad (\text{B.13})$$

$$\{\bar{\tau}_s^{(q)}\}_{q=1}^Q = \bar{\tau}_s^{(j)} \quad (\text{B.14})$$

where the component weights, means, and covariances are computed using the splitting library given in Table 2.1.

$$\bar{w}_s^{(q)} = \tilde{\alpha}_q \bar{w}_s^{(j)} \quad (\text{B.15})$$

$$\bar{\mathbf{x}}_s^{(q)} = \bar{\mathbf{x}}_s^{(j)} + \sqrt{\lambda_r} \tilde{m}_q \mathbf{v}_r \quad (\text{B.16})$$

$$\bar{P}_s^{(q)} = V \Lambda_q V^T \quad (\text{B.17})$$

$$\Lambda_q = \text{diag}\{\lambda_1, \dots, \tilde{\sigma}_q^2 \lambda_r, \dots, \lambda_L\} \quad (\text{B.18})$$

3.D.ii. Compute the sigma points and entropy for each new component using Eqs. (B.5)-(B.6).

3.E. Repeat for all components until time t_k is reached.

3.F. Multiply component weights by the probability of survival.

$$\bar{w}_k^{(j)} = p_{S,k}(\mathbf{x}_k^{(j)}) \bar{w}_k^{(j)} \quad (\text{B.19})$$

3.G. Add process noise to the covariance and recompute sigma points for each component.

$$\bar{P}_k^{(j)} = Q_k + \bar{P}_k^{(j)} \quad (\text{B.20})$$

$$\bar{\chi}_k^{(j)} = \begin{bmatrix} \bar{\mathbf{x}}_k^{(j)} & \bar{\mathbf{x}}_k^{(j)} + \gamma \sqrt{\bar{P}_k^{(j)}} & \bar{\mathbf{x}}_k^{(j)} - \gamma \sqrt{\bar{P}_k^{(j)}} \end{bmatrix}_{L \times (2L+1)} \quad (\text{B.21})$$

There are now \bar{J}_k components.

4. Correction Step

4.A. Missed Detections - Loop through all components $j \in [1, \bar{J}_k]$.

4.A.i. Downweight components based on the probability of detection.

$$w_k^{(j)} = [1 - p_{D,k}(\mathbf{x}^{(j)})] \bar{w}_k^{(j)} \quad (\text{B.22})$$

4.A.ii. Retain *a priori* component information.

$$\hat{\mathbf{x}}_k^{(j)} = \bar{\mathbf{x}}_k^{(j)} \quad (\text{B.23})$$

$$P_k^{(j)} = \bar{P}_k^{(j)} \quad (\text{B.24})$$

$$\chi_k^{(j)} = \bar{\chi}_k^{(j)} \quad (\text{B.25})$$

$$\mathbf{H}_{\text{lin},k}^{(j)} = \bar{\mathbf{H}}_{\text{lin},k}^{(j)} \quad (\text{B.26})$$

$$\tau_k^{(j)} = \bar{\tau}_k^{(j)} \quad (\text{B.27})$$

4.B. Detected Objects

4.B.i. Compute the expected measurement and cross covariance for each component.

$$\zeta_k^{(j)} = G(\bar{\chi}_k^{(j)}, t_k) \quad (\text{B.28})$$

$$\bar{\mathbf{z}}_k^{(j)} = \sum_{l=0}^{2L} W_l^m \zeta_{k,l}^{(j)} \quad (\text{B.29})$$

$$P_{xz}^{(j)} = \sum_{l=0}^{2L} W_l^c (\bar{\chi}_{k,l}^{(j)} - \bar{\mathbf{x}}_k^{(j)}) (\zeta_{k,l}^{(j)} - \bar{\mathbf{z}}_k^{(j)})^T \quad (\text{B.30})$$

4.B.ii. For measurement $\mathbf{z}_k^{(i)} \in Z_k$, $i \in [1, m_k]$, compute the update for each component.

$$P_{zz}^{(i,j)} = R_k^{(i)} + \sum_{l=0}^{2L} W_l^c (\zeta_{k,l}^{(j)} - \bar{\mathbf{z}}_k^{(j)}) (\zeta_{k,l}^{(j)} - \bar{\mathbf{z}}_k^{(j)})^T \quad (\text{B.31})$$

$$K_k^{(i,j)} = P_{xz}^{(j)} [P_{zz}^{(i,j)}]^{-1} \quad (\text{B.32})$$

$$\hat{\mathbf{x}}_k^{(i\bar{J}_k+j)} = \bar{\mathbf{x}}_k^{(j)} + K_k^{(i,j)} (\mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^{(j)}) \quad (\text{B.33})$$

$$P_k^{(i\bar{J}_k+j)} = \bar{P}_k^{(j)} - K_k^{(i,j)} P_{zz}^{(i,j)} [K_k^{(i,j)}]^T \quad (\text{B.34})$$

$$\tau_k^{(i\bar{J}_k+j)} = \bar{\tau}_k^{(j)} \quad (\text{B.35})$$

4.B.iii. Compute the updated sigma points and entropy.

$$\chi_k^{(i\bar{J}_k+j)} = \left[\hat{\mathbf{x}}_k^{(i\bar{J}_k+j)} \quad \hat{\mathbf{x}}_k^{(i\bar{J}_k+j)} + \gamma \sqrt{P_k^{(i\bar{J}_k+j)}} \quad \hat{\mathbf{x}}_k^{(i\bar{J}_k+j)} - \gamma \sqrt{P_k^{(i\bar{J}_k+j)}} \right]_{L \times (2L+1)} \quad (\text{B.36})$$

$$\mathbf{H}_{\text{lin},k}^{(i\bar{J}_k+j)} = \frac{1}{2} \log |2\pi e P_k^{(i\bar{J}_k+j)}| \quad (\text{B.37})$$

4.B.iv. Normalize weights across measurement i and multiply by the probability of detection.

$$g_k^{(i,j)} = p_g(\mathbf{z}_k^{(i)}; \bar{\mathbf{z}}_k^{(j)}, P_{zz}^{(i,j)}) \quad (\text{B.38})$$

$$w_k^{(i\bar{J}_k+j)} = \frac{p_{D,k}(\mathbf{x}_k^{(j)})g_k^{(i,j)}\bar{w}_k^{(j)}}{\kappa_k + \sum_{q=1}^{\bar{J}_k} p_{D,k}(\mathbf{x}_k^{(q)})g_k^{(i,q)}\bar{w}_k^{(q)}} \quad (\text{B.39})$$

There are now $J_{k,0} = (1 + m_k)\bar{J}_k$ components.

5. Merge and Prune GMM Components

5.A. Remove components whose weight is below a threshold. Discard information of components that don't meet the criteria.

$$w_{k,1} = \{w_{k,0}^{(j)} : w_{k,0}^{(j)} \geq T_{\text{prune}} * \max(w_{k,0})\} \quad (\text{B.40})$$

5.B. Renormalize the weights for the current number of components $J_{k,1}$.

$$w_{k,1}^{(j)} = w_{k,0}^{(j)} \frac{\sum_{i_0}^{J_{k,0}} w_{k,0}^{(i_0)}}{\sum_{i_1}^{J_{k,1}} w_{k,1}^{(i_1)}} \quad (\text{B.41})$$

5.C. Merge components that are close together.

5.C.i. Set $I = \{1, \dots, J_{k,1}\}$.

5.C.ii. Repeat until $I = \{\}$:

$$j^* = \operatorname{argmax}(w_{k,1}^{(j)}) \quad (\text{B.42})$$

$$\mathcal{L} = \{i \in I : (\hat{\mathbf{x}}_k^{(i)} - \hat{\mathbf{x}}_k^{(j^*)})^T (P_k^{(j^*)})^{-1} (\hat{\mathbf{x}}_k^{(i)} - \hat{\mathbf{x}}_k^{(j^*)}) \leq T_{\text{merge}}\} \quad (\text{B.43})$$

$$w_k^{(j)} = \sum_{l \in \mathcal{L}} w_{k,1}^{(l)} \quad (\text{B.44})$$

$$\hat{\mathbf{x}}_k^{(j)} = \frac{1}{w_k^{(j)}} \sum_{l \in \mathcal{L}} w_{k,1}^{(l)} \hat{\mathbf{x}}_k^{(l)} \quad (\text{B.45})$$

$$P_k^{(j)} = \frac{1}{w_k^{(j)}} \sum_{l \in \mathcal{L}} w_{k,1}^{(l)} [P_k^{(l)} + (\hat{\mathbf{x}}_k^{(j)} - \hat{\mathbf{x}}_k^{(l)}) (\hat{\mathbf{x}}_k^{(j)} - \hat{\mathbf{x}}_k^{(l)})^T] \quad (\text{B.46})$$

$$\chi_k^{(j)} = \begin{bmatrix} \hat{\mathbf{x}}_k^{(j)} & \hat{\mathbf{x}}_k^{(j)} + \gamma \sqrt{P_k^{(j)}} & \hat{\mathbf{x}}_k^{(j)} - \gamma \sqrt{P_k^{(j)}} \end{bmatrix}_{L \times (2L+1)} \quad (\text{B.47})$$

$$H_{\text{lin},k}^{(j)} = \frac{1}{2} \log |2\pi e P_k^{(j)}| \quad (\text{B.48})$$

$$\tau_k^{(j)} = \tau_k^{(j^*)} \quad (\text{B.49})$$

$$I = I \setminus \mathcal{L} \quad (\text{B.50})$$

5.C.iii. If there are more than J_{max} components, keep only the J_{max} highest weights.

Renormalize weights after pruning.

There are now J_k components.

B.2 AEGIS CPHD Filter

1.A. Initial GMM approximation of PHD at time t_0 with track labels T_0 .

$$\nu_0(\mathbf{x}_0) \approx \sum_{j=1}^{J_0} w_0^{(j)} p_g(\mathbf{x}_0; \hat{\mathbf{x}}_0^{(j)}, P_0^{(j)})$$

$$T_0 = \{\tau_0^{(j)}\}_{j=1}^{J_0}$$

1.B. Initialize cardinality distribution.

1.B.i. Option 1: Uniform cardinality distribution at time t_0 .

$$p_0(n) = 1/(n_{\text{obj,max}} + 1) \quad n \in \{0, \dots, n_{\text{obj,max}}\} \quad (\text{B.51})$$

1.B.ii. Option 2: Multi-Bernoulli cardinality distribution at t_0 .

$$p_0(n) = C_n^{n_{\text{obj,max}}} (1-q)^{n_{\text{obj,max}}} \left(\frac{q}{1-q} \right)^n \quad (\text{B.52})$$

where $q = N/n_{\text{obj,max}}$ and N is the true number of objects.

2. Compute sigma points and entropy for each component.

$$\chi_0^{(j)} = \left[\hat{\mathbf{x}}_0^{(j)} \quad \hat{\mathbf{x}}_0^{(j)} + \gamma \sqrt{P_0^{(j)}} \quad \hat{\mathbf{x}}_0^{(j)} - \gamma \sqrt{P_0^{(j)}} \right]_{L \times (2L+1)} \quad (\text{B.53})$$

$$H_0^{(j)} = \frac{1}{2} \log |2\pi e P_0^{(j)}| \quad (\text{B.54})$$

3. Prediction Step $t \in [t_{k-1}, t_k]$

3.A. Predict cardinality and number of objects (non-spawning case).

$$p_{k|k-1}(n) = \sum_{j=0}^n p_{\Gamma,k}(n-j) \sum_{l=j}^{n_{\text{obj,max}}} C_j^l p_{k-1}(l) p_S^{(j)} (1-p_S)^{(l-j)} \quad (\text{B.55})$$

$$n \in [0, \dots, n_{\text{obj,max}}] \quad (\text{B.56})$$

$$C_j^l = \frac{l!}{j!(l-j)!} \quad (\text{B.57})$$

$$N_{k|k-1} = \sum_{n=1}^{n_{\text{obj,max}}} n \cdot p_{k|k-1}(n) \quad (\text{B.58})$$

3.B. Step through time in increments of $\Delta t_s \leq \Delta t_k$. For $t \in [t_{s-1}, t_s]$, propagate the sigma points and linear PDF entropy for the j -th component to compute $\bar{\chi}_s^{(j)}$ and $\bar{H}_{\text{lin},s}^{(j)}$.

$$\dot{\chi}_{s-1}^{(j)} = F(\chi_{s-1}^{(j)}, t_{s-1}), \quad \chi^{(j)}(t_{s-1}) = \chi_{s-1}^{(j)} \quad (\text{B.59})$$

$$\dot{H}_{\text{lin}}^{(j)} = \text{trace}\{A_{s-1}^{(j)}\}, \quad H_{\text{lin}}^{(j)}(t_{s-1}) = H_{\text{lin},s-1}^{(j)} \quad (\text{B.60})$$

3.C. Compute the component *a priori* mean and covariance, retain track label.

$$\bar{\mathbf{x}}_s^{(j)} = \sum_{l=0}^{2L} W_l^m \bar{\chi}_{s,l}^{(j)} \quad (\text{B.61})$$

$$\bar{P}_s^{(j)} = \sum_{l=0}^{2L} W_l^c (\bar{\chi}_{s,l}^{(j)} - \bar{\mathbf{x}}_s^{(j)}) (\bar{\chi}_{s,l}^{(j)} - \bar{\mathbf{x}}_s^{(j)})^T \quad (\text{B.62})$$

$$\bar{\tau}_s^{(j)} = \tau_{s-1}^{(j)} \quad (\text{B.63})$$

3.D. Compute the entropy from the nonlinear PDF propagation.

$$\bar{H}_{\text{nl},s}^{(j)} = \frac{1}{2} \log |2\pi e \bar{P}_s^{(j)}| \quad (\text{B.64})$$

3.E. Check for nonlinearity. If $|\bar{H}_{\text{nl},s}^{(j)} - \bar{H}_{\text{lin},s}^{(j)}| > T_{\text{split}}$, split along the r -th eigenvector (corresponding to the largest eigenvalue).

3.E.i. Replace the j -th component with Q new components, retain track label of j -th component.

$$\bar{w}_s^{(j)} p_g(\mathbf{x}_s; \bar{\mathbf{x}}_s^{(j)}, \bar{P}_s^{(j)}) \approx \sum_{q=1}^Q \bar{w}_s^{(q)} p_g(\mathbf{x}_s; \bar{\mathbf{x}}_s^{(q)}, \bar{P}_s^{(q)}) \quad (\text{B.65})$$

$$\{\bar{\tau}_s^{(q)}\}_{q=1}^Q = \bar{\tau}_s^{(j)} \quad (\text{B.66})$$

where the component weights, means, and covariances are computed using the splitting library given in Table 2.1.

$$\bar{w}_s^{(q)} = \tilde{\alpha}_q \bar{w}_s^{(j)} \quad (\text{B.67})$$

$$\bar{\mathbf{x}}_s^{(q)} = \bar{\mathbf{x}}_s^{(j)} + \sqrt{\lambda_r} \tilde{m}_q \mathbf{v}_r \quad (\text{B.68})$$

$$\bar{P}_s^{(q)} = V \Lambda_q V^T \quad (\text{B.69})$$

$$\Lambda_q = \text{diag}\{\lambda_1, \dots, \tilde{\sigma}_q^2 \lambda_r, \dots, \lambda_L\} \quad (\text{B.70})$$

3.E.ii. Compute the sigma points and entropy for each new component using Eqs. (B.53)-(B.54).

3.F. Repeat for all components until time t_k is reached.

3.G. Multiply component weights by the probability of survival.

$$\bar{w}_k^{(j)} = p_{S,k}(\mathbf{x}_k^{(j)}) \bar{w}_k^{(j)} \quad (\text{B.71})$$

3.H. Add process noise to the covariance and recompute sigma points for each component.

$$\bar{P}_k^{(j)} = Q_k + \bar{P}_k^{(j)} \quad (\text{B.72})$$

$$\bar{\chi}_k^{(j)} = \begin{bmatrix} \bar{\mathbf{x}}_k^{(j)} & \bar{\mathbf{x}}_k^{(j)} + \gamma \sqrt{\bar{P}_k^{(j)}} & \bar{\mathbf{x}}_k^{(j)} - \gamma \sqrt{\bar{P}_k^{(j)}} \end{bmatrix}_{L \times (2L+1)} \quad (\text{B.73})$$

There are now \bar{J}_k components.

4. Correction Step

4.A. Compute update information for each measurement to component

4.A.i. Compute the expected measurement and cross covariance for each component.

$$\zeta_k^{(j)} = G(\bar{\chi}_k^{(j)}, t_k) \quad (\text{B.74})$$

$$\bar{z}_k^{(j)} = \sum_{l=0}^{2L} W_l^m \zeta_{k,l}^{(j)} \quad (\text{B.75})$$

$$P_{xz}^{(j)} = \sum_{l=0}^{2L} W_l^c (\bar{\chi}_{k,l}^{(j)} - \bar{\mathbf{x}}_k^{(j)}) (\zeta_{k,l}^{(j)} - \bar{z}_k^{(j)})^T \quad (\text{B.76})$$

4.A.ii. For measurement $\mathbf{z}_k^{(i)} \in Z_k$, $i \in [1, m_k]$, compute the measurement to component likelihood.

$$P_{zz}^{(i,j)} = R_k^{(i)} + \sum_{l=0}^{2L} W_l^c (\zeta_{k,l}^{(j)} - \bar{z}_k^{(j)}) (\zeta_{k,l}^{(j)} - \bar{z}_k^{(j)})^T \quad (\text{B.77})$$

$$K_k^{(i,j)} = P_{xz}^{(j)} [P_{zz}^{(i,j)}]^{-1} \quad (\text{B.78})$$

$$g_k^{(i,j)} = p_g(\mathbf{z}_k^{(i)}; \bar{z}_k^{(j)}, P_{zz}^{(i,j)}) \quad (\text{B.79})$$

4.B. Compute likelihood functions

$$\Psi_k^u[w, Z](n) = \sum_{j=0}^{\min(m_k, n)} (m_k - j)! p_\kappa(m_k - j) P_{j+u}^n \frac{\langle 1 - p_D, \nu \rangle^{n-(j+u)}}{\langle 1, \nu \rangle^n} \sigma_j(\{\Lambda_k^{(i)}(w, Z)\}_{i=1}^{m_k}) \quad (\text{B.80})$$

$$\Lambda_k^{(i)}(w, Z) = \left\{ \frac{\langle 1, \kappa_k \rangle}{\kappa_k(z)} p_D \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} g_k^{(i,j)} \right\} \quad (\text{B.81})$$

$$n \in [1, \dots, n_{obj, max}] \quad u = 0, 1$$

An efficient way to compute the $\sigma_j(\cdot)$ coefficients is provided in Reference [49] P. 641.

4.C. Cardinality Update

$$p_k(n) = \frac{\Psi_k^0[w_{k|k-1}, Z_k](n) p_{k|k-1}(n)}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{k|k-1} \rangle} \quad (\text{B.82})$$

$$n \in [1, \dots, n_{obj, max}]$$

4.D. Missed Detections - Loop through all components $j \in [1, \bar{J}_k]$.

4.D.i. Downweight components based on the probability of detection.

$$w_k^{(j)} = \frac{\langle \Psi_k^1[w_{k|k-1}, Z_k], p_{X,k|k-1} \rangle}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{X,k|k-1} \rangle} (1 - p_D) \bar{w}_k^{(j)} \quad (\text{B.83})$$

4.D.ii. Retain *a priori* component information.

$$\hat{\mathbf{x}}_k^{(j)} = \bar{\mathbf{x}}_k^{(j)} \quad (\text{B.84})$$

$$P_k^{(j)} = \bar{P}_k^{(j)} \quad (\text{B.85})$$

$$\chi_k^{(j)} = \bar{\chi}_k^{(j)} \quad (\text{B.86})$$

$$\mathbf{H}_{\text{lin},k}^{(j)} = \bar{\mathbf{H}}_{\text{lin},k}^{(j)} \quad (\text{B.87})$$

$$\tau_k^{(j)} = \bar{\tau}_k^{(j)} \quad (\text{B.88})$$

4.E. Detected Objects

4.E.i. Compute the updated mean and covariance, retain track label.

$$\hat{\mathbf{x}}_k^{(i\bar{J}_k+j)} = \bar{\mathbf{x}}_k^{(j)} + K_k^{(i,j)} (\mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^{(j)}) \quad (\text{B.89})$$

$$P_k^{(i\bar{J}_k+j)} = \bar{P}_k^{(j)} - K_k^{(i,j)} P_{zz}^{(i,j)} [K_k^{(i,j)}]^T \quad (\text{B.90})$$

$$\tau_k^{(i\bar{J}_k+j)} = \bar{\tau}_k^{(j)} \quad (\text{B.91})$$

4.E.ii. Compute the updated sigma points and entropy.

$$\chi_k^{(i\bar{J}_k+j)} = \left[\hat{\mathbf{x}}_k^{(i\bar{J}_k+j)} \quad \hat{\mathbf{x}}_k^{(i\bar{J}_k+j)} + \gamma \sqrt{P_k^{(i\bar{J}_k+j)}} \quad \hat{\mathbf{x}}_k^{(i\bar{J}_k+j)} - \gamma \sqrt{P_k^{(i\bar{J}_k+j)}} \right]_{L \times (2L+1)} \quad (\text{B.92})$$

$$\mathbf{H}_{\text{lin},k}^{(i\bar{J}_k+j)} = \frac{1}{2} \log |2\pi e P_k^{(i\bar{J}_k+j)}| \quad (\text{B.93})$$

4.E.iii. Compute the updated weights

$$w_k^{(i\bar{J}_k+j)} = p_D w_{k|k-1}^{(j)} g_k^{(i,j)} \frac{\langle \Psi_k^1[w_{k|k-1}, Z_k \setminus \{\mathbf{z}^{(i)}\}], p_{X,k|k-1} \rangle}{\langle \Psi_k^0[w_{k|k-1}, Z_k], p_{X,k|k-1} \rangle} \frac{\langle 1, \kappa_k \rangle}{\kappa_k(\mathbf{z}^{(i)})} \quad (\text{B.94})$$

There are now $J_{k,0} = (1 + m_k) \bar{J}_k$ components.

5. Merge and Prune GMM Components

5.A. Remove components whose weight is below a threshold. Discard information of components that don't meet the criteria.

$$w_{k,1}^{(j)} = \{w_{k,0}^{(j)} : w_{k,0}^{(j)} \geq T_{\text{prune}} * \max(w_{k,0}^{(j)})\} \quad (\text{B.95})$$

5.B. Renormalize the weights for the current number of components $J_{k,1}$.

$$w_{k,1}^{(j)} = w_{k,1}^{(j)} \frac{\sum_{i_0}^{J_{k,0}} w_{k,0}^{(i_0)}}{\sum_{i_1}^{J_{k,1}} w_{k,1}^{(i_1)}} \quad (\text{B.96})$$

5.C. Merge components that are close together.

5.C.i. Set $I = \{1, \dots, J_{k,1}\}$.

5.C.ii. Repeat until $I = \{\}$:

$$j^* = \text{argmax}(w_{k,1}^{(j)}) \quad (\text{B.97})$$

$$\mathcal{L} = \{i \in I : (\hat{\mathbf{x}}_k^{(i)} - \hat{\mathbf{x}}_k^{(j^*)})^T (P_k^{(i)})^{-1} (\hat{\mathbf{x}}_k^{(i)} - \hat{\mathbf{x}}_k^{(j^*)}) \leq T_{\text{merge}}\} \quad (\text{B.98})$$

$$w_k^{(j)} = \sum_{l \in \mathcal{L}} w_{k,1}^{(l)} \quad (\text{B.99})$$

$$\hat{\mathbf{x}}_k^{(j)} = \frac{1}{w_k^{(j)}} \sum_{l \in \mathcal{L}} w_{k,1}^{(l)} \hat{\mathbf{x}}_k^{(l)} \quad (\text{B.100})$$

$$P_k^{(j)} = \frac{1}{w_k^{(j)}} \sum_{l \in \mathcal{L}} w_{k,1}^{(l)} [P_k^{(l)} + (\hat{\mathbf{x}}_k^{(j)} - \hat{\mathbf{x}}_k^{(l)})(\hat{\mathbf{x}}_k^{(j)} - \hat{\mathbf{x}}_k^{(l)})^T] \quad (\text{B.101})$$

$$\chi_k^{(j)} = \begin{bmatrix} \hat{\mathbf{x}}_k^{(j)} & \hat{\mathbf{x}}_k^{(j)} + \gamma \sqrt{P_k^{(j)}} & \hat{\mathbf{x}}_k^{(j)} - \gamma \sqrt{P_k^{(j)}} \end{bmatrix}_{L \times (2L+1)} \quad (\text{B.102})$$

$$H_{\text{lin},k}^{(j)} = \frac{1}{2} \log |2\pi e P_k^{(j)}| \quad (\text{B.103})$$

$$\tau_k^{(j)} = \tau_k^{(j^*)} \quad (\text{B.104})$$

$$I = I \setminus \mathcal{L} \quad (\text{B.105})$$

5.C.iii. If there are more than J_{max} components, keep only the J_{max} highest weights.

Renormalize weights after pruning.

There are now J_k components.

Appendix C

Information Gain Equations

C.1 Derivation of Rényi Divergence

Beginning with the general form of the Rényi divergence [66],

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \log \int f_1(X; \mathbf{u})^\alpha f_0(X)^{1-\alpha} dX \quad (\text{C.1})$$

Writing out the set integral in terms of the Janossy densities gives

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \log \sum_{n \geq 0} \frac{1}{n!} \int \left[j_1^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}; \mathbf{u}) \right]^\alpha \left[j_0^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) \right]^{1-\alpha} d\mathbf{x}_1 \dots d\mathbf{x}_n \quad (\text{C.2})$$

In the case of the CPHD filter, the multitarget PDF is modeled as an independent and identically distributed (i.i.d.) cluster process, in which the Janossy density can be represented by a spatial single object density, $s(\mathbf{x})$, and a cardinality distribution. Note that $s(\mathbf{x})$ defines a spatial density for all objects that integrates to one over the single object state space, it is not a PDF representing an individual object.

$$j^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) = n! \cdot p(n) \prod_{i=1}^n s(\mathbf{x}_i) \quad (\text{C.3})$$

Substitution into Eq. (C.2) yields

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \log \sum_{n \geq 0} \frac{1}{n!} \int \left[n! \cdot p_1(n; \mathbf{u}) \prod_{i=1}^n s_1(\mathbf{x}_i; \mathbf{u}) \right]^\alpha \left[n! \cdot p_0(n) \prod_{i=1}^n s_0(\mathbf{x}_i) \right]^{1-\alpha} d\mathbf{x}_i \quad (\text{C.4})$$

Canceling the $n!$ and rearranging terms produces

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \log \sum_{n \geq 0} p_1(n; \mathbf{u})^\alpha p_0(n)^{1-\alpha} \prod_{i=1}^n \int s_1(\mathbf{x}_i; \mathbf{u})^\alpha s_0(\mathbf{x}_i)^{1-\alpha} d\mathbf{x}_i \quad (\text{C.5})$$

Each term in the product is equal, allowing it to be replaced by an exponent to produce the final form of the CPHD Rényi divergence as given in Eq. (14) of Reference [66].

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \log \sum_{n \geq 0} p_1(n; \mathbf{u})^\alpha p_0(n)^{1-\alpha} \left[\int s_1(\mathbf{x}; \mathbf{u})^\alpha s_0(\mathbf{x})^{1-\alpha} d\mathbf{x} \right]^n \quad (\text{C.6})$$

To apply the GMM approximation, note that the density $s(\mathbf{x})$ is related to the PHD by

$$\nu(\mathbf{x}) = s(\mathbf{x}) \sum_{n=1}^{\infty} n \cdot p(n) \quad (\text{C.7})$$

This allows the Rényi divergence to be written in terms of the cardinality and PHD function,

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \log \sum_{n \geq 0} \left(\frac{p_1(n; \mathbf{u})}{N_1^n} \right)^\alpha \left(\frac{p_0(n)}{N_0^n} \right)^{1-\alpha} \left[\int \nu_1(\mathbf{x}; \mathbf{u})^\alpha \nu_0(\mathbf{x})^{1-\alpha} d\mathbf{x} \right]^n \quad (\text{C.8})$$

where N_k is the expected target number. Note that in Ch. 3 a reduced cardinality distribution is computed for components in the field of view. The weights of the components are truncated to a maximum value of 0.999 in order to compute the multi-Bernoulli distribution for the cardinality. In order for the equation above to normalize correctly, the expected target number must be computed using the true component weights, $N_k = \sum_{j=1}^{J_{\text{FOV}}} w_j$, as opposed to using the truncated weights or cardinality within the FOV.

The choice of $\alpha = 0.5$ provides the best discrimination between PDFs [66, 30]. With this selection and assuming a Gaussian mixture approximation of the PHD, the final equation is given by

$$\nu_0(\mathbf{x}) \approx \sum_{i=1}^{J_0} w_i p_g(\mathbf{x}; \mathbf{m}_i, P_i) \quad \nu_1(\mathbf{x}; \mathbf{u}) \approx \sum_{j=1}^{J_1} w_j p_g(\mathbf{x}; \mathbf{m}_j, P_j) \quad (\text{C.9})$$

$$R(\mathbf{u}) \approx -2 \log \sum_{n \geq 0} \left(\frac{p_1(n; \mathbf{u})}{N_1^n} \right)^{1/2} \left(\frac{p_0(n)}{N_0^n} \right)^{1/2} \cdot \left[\int \left(\sum_{i=1}^{J_0} \sum_{j=1}^{J_1} w_i w_j K_{i,j} p_g(\mathbf{x}; \mathbf{m}_{i,j}, P_{i,j}) \right)^{1/2} d\mathbf{x} \right]^n \quad (\text{C.10})$$

$$\begin{aligned} K_{i,j} &= p_g(\mathbf{m}_i; \mathbf{m}_j, P_i + P_j) \\ P_{i,j} &= \left[P_i^{-1} + P_j^{-1} \right]^{-1} \\ \mathbf{m}_{i,j} &= P_{i,j} \left[P_i^{-1} \mathbf{m}_i + P_j^{-1} \mathbf{m}_j \right] \end{aligned}$$

Derivation of the PHD form is straightforward. The PHD filter assumes a Poisson distributed number of targets, replacing the cardinality with $p(n) = \frac{e^{-N} N^n}{n!}$. Substituting in Eq. (C.8) gives

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \log \sum_{n \geq 0} \left(\frac{e^{-N_1}}{n!} \right)^\alpha \left(\frac{e^{-N_0}}{n!} \right)^{1-\alpha} \left[\int \nu_1(\mathbf{x}; \mathbf{u})^\alpha \nu_0(\mathbf{x})^{1-\alpha} d\mathbf{x} \right]^n \quad (\text{C.11})$$

After rearranging terms, the expression can be simplified by making use of the identity $e^x = \sum_{n \geq 0} \frac{x^n}{n!}$ and computing the logarithm.

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \log e^{-N_1 \alpha} e^{-N_0(1-\alpha)} \sum_{n \geq 0} \frac{1}{n!} \left[\int \nu_1(\mathbf{x}; \mathbf{u})^\alpha \nu_0(\mathbf{x})^{1-\alpha} d\mathbf{x} \right]^n \quad (\text{C.12})$$

$$R(\mathbf{u}) = \frac{1}{\alpha - 1} \left[-N_1 \alpha - N_0(1 - \alpha) + \int \nu_1(\mathbf{x}; \mathbf{u})^\alpha \nu_0(\mathbf{x})^{1-\alpha} d\mathbf{x} \right] \quad (\text{C.13})$$

This yields the final form of the PHD Rényi divergence, as given in Eq. (18) of Reference [66].

$$R(\mathbf{u}) = N_0 + \frac{\alpha}{1 - \alpha} N_1 + \frac{1}{\alpha - 1} \int \nu_1(\mathbf{x}; \mathbf{u})^\alpha \nu_0(\mathbf{x})^{1-\alpha} d\mathbf{x} \quad (\text{C.14})$$

Making use of the GMM approximation and setting $\alpha = 0.5$ yields

$$R(\mathbf{u}) \approx \sum_{i=1}^{J_0} w_i + \sum_{j=1}^{J_1} w_j - 2 \int \left(\sum_{i=1}^{J_0} \sum_{j=1}^{J_1} w_i w_j K_{i,j} p_g(\mathbf{x}; \mathbf{m}_{i,j}, P_{i,j}) \right)^{1/2} d\mathbf{x} \quad (\text{C.15})$$

C.2 Simplification of Single Target Case

In the case that a single target is present, the equations above can be simplified to produce an analytic result. Both the PHD and CPHD forms require computing the following integral,

$$I = \int [w_0 w_1 K_{0,1} \cdot p_g(\mathbf{x}; \mathbf{m}_{0,1}, P_{0,1})]^{1/2} d\mathbf{x} \quad (\text{C.16})$$

where the subscripts are used to indicate quantities from the prior and posterior distributions as opposed to GMM component indices. Using the PIMS approximation, the mean $\mathbf{m}_1 = \mathbf{m}_0$, and the following simplification can be made.

$$K_{0,1} = \frac{1}{\sqrt{|2\pi(P_0 + P_1)|}} \quad (\text{C.17})$$

Rearranging terms and writing out the Gaussian PDF explicitly gives

$$I = \left(\frac{w_0 \cdot w_1}{\sqrt{|2\pi(P_0 + P_1)|}} \right)^{1/2} \int \left(\frac{1}{\sqrt{|2\pi P_{0,1}|}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mathbf{m}_{0,1})^T P_{0,1}^{-1} (\mathbf{x} - \mathbf{m}_{0,1}) \right] \right)^{1/2} d\mathbf{x} \quad (\text{C.18})$$

$$I = \left(\frac{w_0 \cdot w_1}{\sqrt{|2\pi(P_0 + P_1)|}} \right)^{1/2} \left(\frac{1}{\sqrt{|2\pi P_{0,1}|}} \right)^{1/2} \int \exp \left[-\frac{1}{4} (\mathbf{x} - \mathbf{m}_{0,1})^T P_{0,1}^{-1} (\mathbf{x} - \mathbf{m}_{0,1}) \right] d\mathbf{x} \quad (\text{C.19})$$

The remaining integral has an analytic solution.

$$\int \exp \left[-\frac{1}{4} (\mathbf{x} - \mathbf{m}_{0,1})^T P_{0,1}^{-1} (\mathbf{x} - \mathbf{m}_{0,1}) \right] d\mathbf{x} = |4\pi P_{0,1}|^{1/2} \quad (\text{C.20})$$

Substituting and rearranging terms produces the final value of the integral.

$$I = \left(\frac{w_0 \cdot w_1}{\sqrt{|2\pi(P_0 + P_1)|}} \right)^{1/2} \left(\frac{1}{\sqrt{|2\pi P_{0,1}|}} \right)^{1/2} |4\pi P_{0,1}|^{1/2} \quad (\text{C.21})$$

$$I = \sqrt{w_0 w_1} \left(\frac{|4\pi P_{0,1}|^{1/4} \cdot |4\pi P_{0,1}|^{1/4}}{|2\pi(P_0 + P_1)|^{1/4} \cdot |2\pi P_{0,1}|^{1/4}} \right) \quad (\text{C.22})$$

$$I = \sqrt{w_0 w_1} \left(\frac{|4\pi P_{0,1}|^{1/4}}{|\pi(P_0 + P_1)|^{1/4}} \right) \quad (\text{C.23})$$

$$I = \sqrt{w_0 w_1} \left(\frac{|4P_{0,1}|}{|P_0 + P_1|} \right)^{1/4} \quad (\text{C.24})$$

This leads to the final form of the single target equations for Rényi divergence, where the PHD has been approximated using a single Gaussian PDF.

$$R_C \approx -2 \log \sum_{n \geq 0} \left(\frac{p_1(n)}{N_1^n} \right)^{1/2} \left(\frac{p_0(n)}{N_0^n} \right)^{1/2} (w_0 w_1)^{n/2} \left(\frac{|4P_{0,1}|}{|P_0 + P_1|} \right)^{n/4} \quad (\text{C.25})$$

$$R_P \approx w_0 + w_1 - 2\sqrt{w_0 w_1} \left(\frac{|4P_{0,1}|}{|P_0 + P_1|} \right)^{1/4} \quad (\text{C.26})$$

In the CPHD equation, it is necessary to compute the expected target number N_k using the true weights or cardinality distribution in the case that w_0 and w_1 are not unity, i.e., do not assume $N_k = 1$.

C.3 Computation of Numeric Integral

In the case that multiple targets are present, the integral must be computed numerically. An appropriate technique is Monte Carlo integration, using importance sampling to reduce the number

of points needed for a given level of accuracy. The integral is computed from

$$I = \frac{1}{N_{\text{MC}}} \sum_{m=1}^{N_{\text{MC}}} \frac{f(\mathbf{x}_m)}{p(\mathbf{x}_m)} \quad (\text{C.27})$$

where $f(\mathbf{x})$ is the function to be integrated, $p(\mathbf{x})$ is the sampling function, and N_{MC} is the total number of points sampled. For the GMM approximated Rényi divergence, the integrand is

$$f(\mathbf{x}) = \left(\sum_{i=1}^{J_0} \sum_{j=1}^{J_1} w_i w_j K_{i,j} p_g(\mathbf{x}; \mathbf{m}_{i,j}, P_{i,j}) \right)^{1/2} \quad (\text{C.28})$$

The sampling function is chosen to produce a large number of points that contribute significantly to the integral while precluding points that make little or no contribution. Choosing an appropriate function is important to reduce the number of points needed, but any choice will produce a correct result provided there are a sufficient number of points in every region where $f(\mathbf{x}) > 0$. To compute the Rényi divergence, the posterior PHD provides a good option, as it guarantees a large number of points around the peaks of the function. Notably, it is a better choice than the prior PHD which is more spread out. Points sampled from the prior PHD in regions where the posterior PHD is near zero will not contribute to the integral of the product. The function must be normalized such that $\int p(\mathbf{x}) = 1$, yielding two similar alternatives.

$$p(\mathbf{x}) = \frac{1}{J_1} \sum_{j=1}^{J_1} p_g(\mathbf{x}; \mathbf{m}_j, P_j) \quad (\text{C.29})$$

$$p(\mathbf{x}) = \frac{1}{N_1} \sum_{j=1}^{J_1} w_j p_g(\mathbf{x}; \mathbf{m}_j, P_j) \quad (\text{C.30})$$

where $N_1 = \sum_{j=1}^{J_1} w_j$. The results presented in Ch. 3 are computed using Eq. (C.29) with a fixed number of points for each GMM component, though it should be noted that Eq. (C.30) will generally require fewer points by reducing the number needed for components where $w_j < 1$. This feature prompted the use of Eq. (C.30) to compute information gain in Ch. 4, which involved a large number of low weight components as a result of the CAR IOD method. In testing, both methods produced the same result, as expected.

In order to determine the appropriate number of points, the results of numerical integration are compared against the analytic solution for a single object case. One hundred trials are

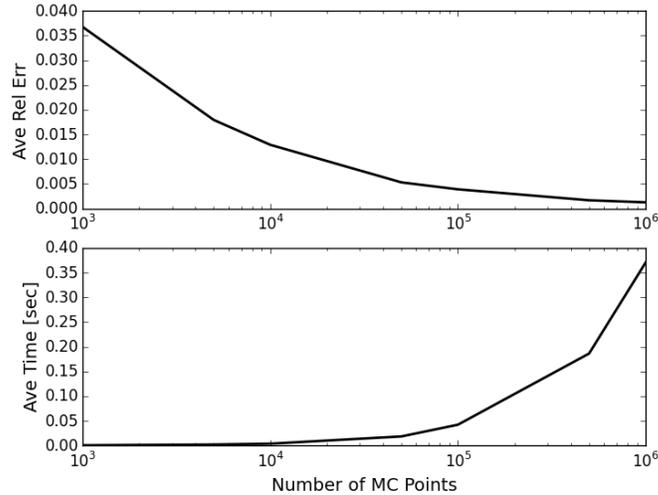


Figure C.1: Numerical Integration Test Case

conducted, in which the prior and posterior covariances are randomly generated, and the average relative errors and execution times are plotted as a function of the number of points used. Figure C.1 provides the results of the test case. The choice of $N_{MC} = 10^5$ produces an average relative error of approximately 0.5%, which is sufficiently accurate for the sensor tasking reward function. It also requires significantly less time than using $N_{MC} = 10^6$ and is therefore selected for all test cases presented in this dissertation.

C.4 Interpreting Units of Information Gain

The output of computing Rényi divergence is a number given in units of nats, the conceptual equivalent of bits but for base e logarithms instead of base 2. The objective function accounts for changes in both the cardinality and PHD function, but does not readily admit a simple physical interpretation for how much information is contained in a given measurement. The following analysis is included to demonstrate, at least to the order of magnitude, the information gain expected for a simplified update to the multitarget PDF.

Assume a single object is present, represented by a single GMM component. Assume the update produces $w_0 = w_1 = 1$, $\mathbf{m}_0 = \mathbf{m}_1$, and that the updated covariance can be defined as a

scalar multiple, $0 < r \leq 1$, of the original, such that

$$P_1 = rP_0 \quad (\text{C.31})$$

In this case, the ratio

$$\frac{|4P_{0,1}|}{|P_0 + P_1|} = \left(\frac{(4r)^p}{(r+1)^{2p}} \right) \quad (\text{C.32})$$

where p is the dimension of the covariance matrix. If there are no changes in the cardinality, the expected information gain becomes

$$R_C \approx -2 \log \left(\frac{(4r)^p}{(r+1)^{2p}} \right)^{1/4} \quad (\text{C.33})$$

For the 2D angles-only measurement, the information gain can therefore be computed as a function of the ratio r , given in Figure C.2(a). From the figure, one natural unit of information gain corresponds to approximately a factor of 10 reduction in the covariance.

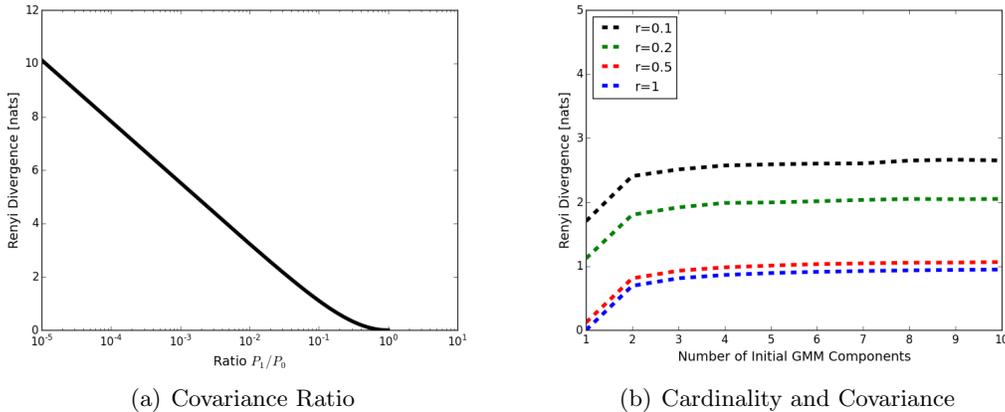


Figure C.2: Rényi Divergence

This is an incomplete analysis, however, as changes in the cardinality also affect the expected information gain. On the right, Figure C.2(b) provides the expected information gain for the case where the initial GMM contains multiple components, which are merged into a single component as part of the update. The cardinality for both the prior and posterior distributions is taken to be a multi-Bernoulli distribution, computed from the component weights, which are evenly distributed, i.e. $w_j = 1/J$. The final covariance is computed by merging the original components

and multiplying by the ratio r . The figure includes results for several values of the covariance reduction ratio.

The blue curve shows the effect of merging components and changing the cardinality without producing any changes in the covariance. For the case of a single initial component and no change in the covariance, no information is gained, as expected. However, even for a small change in the cardinality resulting from a reduction from 2 components to 1, the information gain increases to nearly one natural unit. For the case of a factor of 10 reduction in the covariance, a change in the number of components and associated cardinality can produce an expected information gain greater than 2 nats.

Neither of these tests demonstrate the effects of a realistic measurement update on the expected information gain calculation, however they do provide a rough order of magnitude analysis regarding the changes in cardinality and covariance that correspond to a natural unit of information gain. For a more thorough analysis of the expected information gain due to simulated measurements, please refer to the results of Chapter 4.

Appendix D

Probabilities for the CPHD Filter Birth Model

When implementing a measurement-based birth model for the CPHD filter, an important consideration is the amount of probability mass to associate with new targets as they are introduced to the filter. This appendix seeks to develop an approach to approximate the probability that a measurement represents a new target, which is then used to weight the CAR GMM incorporated in the birth model. The ratio of Gaussian likelihoods used for existing targets and the GMM clutter model is similar to the probability functions developed in Reference [88] for occluded targets.

At time t_k , the multitarget measurement set is represented by the RFS $Z_k = \{\mathbf{z}_1, \dots, \mathbf{z}_{m_k}\}$. Each measurement \mathbf{z}_i is assumed to represent at most one target, and must therefore originate from an existing object, a new object, or a false alarm, producing the following constraint on the probabilities of each.

$$p_{\text{exist}}(\mathbf{z}_i) + p_{\text{new}}(\mathbf{z}_i) + p_{\text{FA}}(\mathbf{z}_i) = 1 \quad (\text{D.1})$$

D.1 Computation of False Alarm Probability

Given the measurement set Z_k and the cardinality and PHD of clutter, it is possible to compute a probability that \mathbf{z}_i represents a false alarm.

$$p_{\text{FA}}(\mathbf{z}_i|Z_k) = \sum_{m=1}^{m_k} p_{\kappa}(m) \cdot p_{\text{FA}}(\mathbf{z}_i|\frac{m}{m_k}) \quad (\text{D.2})$$

where $p_{\kappa}(\cdot)$ is the clutter cardinality and $p_{\text{FA}}(\mathbf{z}_i|\frac{m}{m_k})$ is the probability that \mathbf{z}_i is a false alarm given that there are m false alarms out of m_k measurements, as computed from the clutter PHD. Note

that by definition

$$p_{\text{FA}}(\mathbf{z}_i | \frac{0}{m_k}) = 0 \quad p_{\text{FA}}(\mathbf{z}_i | \frac{m_k}{m_k}) = 1 \quad (\text{D.3})$$

as they respectively represent the cases of zero and all false alarms. The probability that $m > m_k$ measurements are false alarms is zero, producing the upper limit on the summation above.

D.1.1 Poisson Clutter Model

The standard clutter model assumes a Poisson number of false alarms distributed uniformly in the field of view. The clutter cardinality and PHD are defined by

$$p_{\kappa}(m) = \frac{\lambda_{\kappa}^m}{m!} \exp(-\lambda_{\kappa}) \quad (\text{D.4})$$

$$\kappa(\mathbf{z}) = \lambda_{\kappa} \cdot \mathcal{U}(\mathbf{z}) \quad (\text{D.5})$$

$$\mathcal{U}(\mathbf{z}) = \begin{cases} 1/V_s & \text{if } \mathbf{z} \in \text{FOV} \\ 0 & \text{if } \mathbf{z} \notin \text{FOV} \end{cases} \quad (\text{D.6})$$

where V_s is the sensor volume and λ_{κ} is the mean number of clutter returns each scan. Given m_k measurements, there are $C_m^{m_k}$ subsets of m false alarms, each of which is equally likely. The false alarm probability is determined from the sum of the subsets that contain \mathbf{z}_i ,

$$p_{\text{FA}}(\mathbf{z}_i | \frac{m}{m_k}) = \frac{1}{C_m^{m_k}} \sum_{\substack{\pi \subseteq Z_k: \\ |\pi|=m}} \mathbf{1}(\mathbf{z}_i) \quad (\text{D.7})$$

where $\mathbf{1}(\mathbf{z}_i)$ is the indicator function

$$\mathbf{1}(\mathbf{z}_i) = \begin{cases} 1 & \text{if } \mathbf{z}_i \in \pi \\ 0 & \text{if } \mathbf{z}_i \notin \pi \end{cases} \quad (\text{D.8})$$

There are $C_{m-1}^{m_k-1}$ subsets of dimension m that contain \mathbf{z}_i . The probability is therefore given by,

$$p_{\text{FA}}(\mathbf{z}_i | \frac{m}{m_k}) = \frac{C_{m-1}^{m_k-1}}{C_m^{m_k}} = \frac{m}{m_k} \quad (\text{D.9})$$

which yields the full probability that \mathbf{z}_i is a false alarm.

$$p_{\text{FA}}(\mathbf{z}_i | Z_k) = \sum_{m=1}^{m_k} p_{\kappa}(m) \cdot \frac{m}{m_k} \quad (\text{D.10})$$

D.1.2 GMM Clutter Model

Alternately, the clutter PHD can be represented by a GMM, including the probability of detection.

$$\kappa(\mathbf{z}) = \sum_{j=1}^J p_D^{(j)} w_j p_g(\mathbf{z}; \bar{\mathbf{z}}_j, P_{zz}^{(j)}) \quad (\text{D.11})$$

Applying the state-dependent probability of detection as described in Ch. 3 retains only terms within the FOV, all of which are detected with the same constant value p_D .

$$\kappa(\mathbf{z}) = p_D \sum_{j=1}^{J_{\text{FOV}}} w_j p_g(\mathbf{z}; \bar{\mathbf{z}}_j, P_{zz}^{(j)}) \quad (\text{D.12})$$

Similarly, the cardinality can be computed as a multi-Bernoulli distribution using components in the FOV.

$$p_\kappa(m) = \left(\prod_{j=1}^{J_{\text{FOV}}} (1 - p_D w_j) \right) \cdot \sigma_m \left(\left\{ \frac{p_D w_1}{1 - p_D w_1}, \dots, \frac{p_D w_{J_{\text{FOV}}}}{1 - p_D w_{J_{\text{FOV}}}} \right\} \right) \quad (\text{D.13})$$

The probability that measurement \mathbf{z}_i is a false alarm is computed by considering the measurement likelihood, and computing a ratio from the measurement-to-component associations that include \mathbf{z}_i over all possible measurement-to-component associations. As an example, consider the case $m = 1$ false alarm, and apply the state-dependent probability of detection calculation.

$$p_{\text{FA}}(\mathbf{z}_i | \frac{1}{m_k}) = \frac{\sum_{j=1}^{J_{\text{FOV}}} p_D w_j p_g(\mathbf{z}_i; \bar{\mathbf{z}}_j, P_{zz}^{(j)})}{\sum_{l=1}^{m_k} \sum_{j=1}^{J_{\text{FOV}}} p_D w_j p_g(\mathbf{z}_l; \bar{\mathbf{z}}_j, P_{zz}^{(j)})} \quad (\text{D.14})$$

Note that the constant p_D term will cancel, and is therefore omitted going forward. For the case $m = 2$, the number of successful trials are the choices of 2 out of m_k measurements that contain \mathbf{z}_i . The abbreviation $g_{ij} = p_g(\mathbf{z}_i; \bar{\mathbf{z}}_j, P_{zz}^{(j)})$ is adopted for the measurement likelihood.

$$p_{\text{FA}}(\mathbf{z}_i | \frac{2}{m_k}) = \frac{\sum_{\substack{l=1 \\ l \neq i}}^{m_k} \sum_{j=1}^{J_{\text{FOV}}} (w_j g_{ij} + w_j g_{lj})}{\sum_{q=1}^{m_k} \sum_{\substack{l=1 \\ l \neq q}}^{m_k} \sum_{j=1}^{J_{\text{FOV}}} (w_j g_{qj} + w_j g_{lj})} \quad (\text{D.15})$$

The terms in the above equation can be regrouped. There are $(m_k - 1)!$ measurement pairs that contain the term g_{ij} , while the remaining $g_{lj} : l \neq i$ terms only appear $(m_k - 2)!$ times in the

numerator. By extension of the first statement, each g_{lj} term appears $(m_k - 1)!$ times in the denominator.

$$p_{\text{FA}}(\mathbf{z}_i | \frac{2}{Z_k}) = \frac{(m_k - 1)! \sum_{j=1}^{J_{\text{FOV}}} w_j g_{ij} + (m_k - 2)! \sum_{\substack{l=1 \\ l \neq i}}^{m_k} \sum_{j=1}^{J_{\text{FOV}}} w_j g_{lj}}{(m_k - 1)! \sum_{l=1}^{m_k} \sum_{j=1}^{J_{\text{FOV}}} w_j g_{lj}} \quad (\text{D.16})$$

The equation can now be generalized for m false alarms.

$$p_{\text{FA}}(\mathbf{z}_i | \frac{m}{m_k}) = \frac{C_{m-1}^{m_k-1} \sum_{j=1}^{J_{\text{FOV}}} w_j g_{ij} + C_{m-2}^{m_k-2} \sum_{\substack{l=1 \\ l \neq i}}^{m_k} \sum_{j=1}^{J_{\text{FOV}}} w_j g_{lj}}{C_{m-1}^{m_k-1} \sum_{l=1}^{m_k} \sum_{j=1}^{J_{\text{FOV}}} w_j g_{lj}} \quad (\text{D.17})$$

The binomial coefficients can be canceled to produce the final form of the equation, written with the measurement likelihoods included explicitly.

$$p_{\text{FA}}(\mathbf{z}_i | \frac{m}{m_k}) = \frac{\sum_{j=1}^{J_{\text{FOV}}} w_j p_g(\mathbf{z}_i; \bar{\mathbf{z}}_j, P_{zz}^{(j)}) + \frac{m-1}{m_k-1} \sum_{\substack{l=1 \\ l \neq i}}^{m_k} \sum_{j=1}^{J_{\text{FOV}}} w_j p_g(\mathbf{z}_l; \bar{\mathbf{z}}_j, P_{zz}^{(j)})}{\sum_{l=1}^{m_k} \sum_{j=1}^{J_{\text{FOV}}} w_j p_g(\mathbf{z}_l; \bar{\mathbf{z}}_j, P_{zz}^{(j)})} \quad (\text{D.18})$$

In any case where $m < 2$, the term $\frac{m-1}{m_k-1}$ is taken to be zero as a result of the definition of the original binomial coefficient, see [50], Eq. 2.1.

$$C_i^n = \begin{cases} \frac{n!}{i!(n-i)!} & \text{if } 0 \leq i \leq n \\ 0 & \text{if otherwise} \end{cases} \quad (\text{D.19})$$

The final probability for the GMM clutter model is computed as before.

$$p_{\text{FA}}(\mathbf{z}_i | Z_k) = \sum_{m=1}^{m_k} p_{\kappa}(m) \cdot p_{\text{FA}}(\mathbf{z}_i | \frac{m}{m_k}) \quad (\text{D.20})$$

D.1.3 Combined Clutter Model

In the case that both the Poisson and GMM clutter models are used, the clutter PHD is given by the sum.

$$\kappa(\mathbf{z}) = \lambda_{\kappa} \cdot \mathcal{U}(\mathbf{z}) + p_D \sum_{j=1}^{J_{\text{FOV}}} w_j p_g(\mathbf{z}; \bar{\mathbf{z}}_j, P_{zz}^{(j)}) \quad (\text{D.21})$$

The combined cardinality is the convolution of the Poisson and GMM cardinalities,

$$p_{\kappa}(m) = (p_{\kappa,p} * p_{\kappa,g})(m) = \sum_{i=0}^m p_{\kappa,p}(i)p_{\kappa,g}(m-i) \quad (\text{D.22})$$

where the subscripts p and g denote the Poisson and GMM model respectively. The probability $p_{\text{FA}}(\mathbf{z}_i | \frac{m}{m_k})$ is computed from the individual contributions of the Poisson and GMM clutter. In the case that $m = 1$ false alarm is present, it may be due to either the Poisson or GMM model.

$$\begin{aligned} p_{\text{FA}}(\mathbf{z}_i | \frac{1}{m_k}) &= \frac{p_{\kappa,p}(1)p_{\kappa,g}(0)}{p_{\kappa,p}(1)p_{\kappa,g}(0) + p_{\kappa,p}(0)p_{\kappa,g}(1)} p_{\text{FA,p}}(\mathbf{z}_i | \frac{1}{m_k}) \\ &+ \frac{p_{\kappa,p}(0)p_{\kappa,g}(1)}{p_{\kappa,p}(1)p_{\kappa,g}(0) + p_{\kappa,p}(0)p_{\kappa,g}(1)} p_{\text{FA,g}}(\mathbf{z}_i | \frac{1}{m_k}) \end{aligned} \quad (\text{D.23})$$

To develop the more general form, consider the case of 3 measurements, $Z_k = \{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\}$, 2 of which are false alarms.

$$\begin{aligned} p_{\text{FA}}(\mathbf{z}_1 | \frac{2}{m_k}) &= \frac{p_{\kappa,p}(2)p_{\kappa,g}(0)}{\sum_{l=0}^2 p_{\kappa,p}(l)p_{\kappa,g}(2-l)} p_{\text{FA,p}}(\mathbf{z}_1 | \frac{2}{m_k}) \\ &+ \frac{p_{\kappa,p}(1)p_{\kappa,g}(1)}{\sum_{l=0}^2 p_{\kappa,p}(l)p_{\kappa,g}(2-l)} p_{\text{FA,p,g}}(\mathbf{z}_1 | \frac{2}{m_k}) \\ &+ \frac{p_{\kappa,p}(0)p_{\kappa,g}(2)}{\sum_{l=0}^2 p_{\kappa,p}(l)p_{\kappa,g}(2-l)} p_{\text{FA,g}}(\mathbf{z}_1 | \frac{2}{m_k}) \end{aligned} \quad (\text{D.24})$$

The middle term accounts for the possibility of 1 false alarm due to each model, and requires further examination. Given that there is 1 Poisson false alarm, there is an equal probability that it is \mathbf{z}_1 , \mathbf{z}_2 , or \mathbf{z}_3 . In the case that \mathbf{z}_1 is the Poisson false alarm, $p_{\text{FA}}(\mathbf{z}_1) = 1$. In the case that one of the other measurements is the Poisson false alarm, $p_{\text{FA}}(\mathbf{z}_1)$ is computed from the GMM model, with the Poisson false alarm removed from the measurement set. The probability is therefore given by

$$p_{\text{FA,p,g}}(\mathbf{z}_1 | \frac{2}{m_k}) = \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot p_{\text{FA,g}}\left(\mathbf{z}_1 | \frac{1}{|Z_k \setminus \{\mathbf{z}_2\}|}\right) + \frac{1}{3} \cdot p_{\text{FA,g}}\left(\mathbf{z}_1 | \frac{1}{|Z_k \setminus \{\mathbf{z}_3\}|}\right) \quad (\text{D.25})$$

The general form is

$$p_{\text{FA}}(\mathbf{z}_i | \frac{m}{m_k}) = \sum_{j=0}^m \frac{p_{\kappa,p}(j)p_{\kappa,g}(m-j)}{\sum_{l=0}^m p_{\kappa,p}(l)p_{\kappa,g}(m-l)} \left[\frac{1}{C_j^{m_k}} \sum_{\substack{\pi \subseteq Z_k: \\ |\pi|=j}} \left(\mathbf{1}(\mathbf{z}_i) + p_{\text{FA,g}}(\mathbf{z}_i | \frac{m-j}{|Z_k \setminus \pi|}) \right) \right] \quad (\text{D.26})$$

where the convention $p_{\text{FA}}(\mathbf{z}_i|\frac{0}{0}) = 0$ is applied. Using Eq. (D.9), the equation may be rewritten equivalently as

$$p_{\text{FA}}(\mathbf{z}_i|\frac{m}{m_k}) = \sum_{j=0}^m \frac{p_{\kappa,p}(j)p_{\kappa,g}(m-j)}{\sum_{l=0}^m p_{\kappa,p}(l)p_{\kappa,g}(m-l)} \left[\frac{j}{m_k} + \frac{1}{C_j^{m_k}} \sum_{\substack{\pi \subseteq Z_k: \\ |\pi|=j}} p_{\text{FA,g}}(\mathbf{z}_i|\frac{m-j}{|Z_k \setminus \pi|}) \right] \quad (\text{D.27})$$

The final probability for the combined clutter model is computed as before.

$$p_{\text{FA}}(\mathbf{z}_i|Z_k) = \sum_{m=1}^{m_k} p_{\kappa}(m) \cdot p_{\text{FA}}(\mathbf{z}_i|\frac{m}{Z_k}) \quad (\text{D.28})$$

D.2 Computation of New Target Probability

The previous section can be extended to compute the probability that a measurement represents either a false alarm or an existing target by including the GMM model of existing targets with the clutter model. As before, the cardinality of existing targets is computed as a multi-Bernoulli distribution over components in the FOV.

$$p_e(m) = \left(\prod_{j=1}^{J_{\text{FOV}}} (1 - p_D w_j) \right) \cdot \sigma_m \left(\left\{ \frac{p_D w_1}{1 - p_D w_1}, \dots, \frac{p_D w_{J_{\text{FOV}}}}{1 - p_D w_{J_{\text{FOV}}}} \right\} \right) \quad (\text{D.29})$$

The combined GMM cardinality is the convolution of the clutter and existing GMM cardinalities,

$$p_{\kappa g,e}(m) = (p_{\kappa,g} * p_e)(m) = \sum_{i=0}^m p_{\kappa,g}(i) p_e(m-i) \quad (\text{D.30})$$

The conditional probability is computed from Eq. (D.27), with the GMM model including all clutter and existing components in the FOV.

$$p_{\text{FAE}}(\mathbf{z}_i|\frac{m}{m_k}) = \sum_{j=0}^m \frac{p_{\kappa,p}(j)p_{\kappa g,e}(m-j)}{\sum_{l=0}^m p_{\kappa,p}(l)p_{\kappa g,e}(m-l)} \left[\frac{j}{m_k} + \frac{1}{C_j^{m_k}} \sum_{\substack{\pi \subseteq Z_k: \\ |\pi|=j}} p_{\text{FA,ge}}(\mathbf{z}_i|\frac{m-j}{|Z_k \setminus \pi|}) \right] \quad (\text{D.31})$$

where $p_{\text{FAE}}(\mathbf{z}_i|\frac{m}{m_k})$ denotes the probability that a measurement originates from a false alarm or an existing target. The final probability is computed from the sum.

$$p_{\text{FAE}}(\mathbf{z}_i|Z_k) = \sum_{m=1}^{m_k} p_{\kappa p,g,e}(m) \cdot p_{\text{FAE}}(\mathbf{z}_i|\frac{m}{m_k}) \quad (\text{D.32})$$

The probability that \mathbf{z}_i represents a new target can be determined from the constraint in Eq. (D.1).

$$p_{\text{new}}(\mathbf{z}_i|Z_k) = 1 - p_{\text{FAE}}(\mathbf{z}_i|Z_k) \quad (\text{D.33})$$

Appendix E

Spawning Model for the CPHD Filter

This appendix presents the development of a simple CPHD spawning model that can be used to initiate tracks for new objects in the filter. Spawning models assume that new targets originate from known targets, and therefore provide a useful alternative to the target birth model discussed in Chapter 4 because they do not require a solution to the IOD problem. Instead, tracks for new objects are initialized using the estimated state of the original object, with a small deviation added to the state. The spawning model is also highly applicable to the orbital debris problem, as many new objects result from satellite breakups, collisions, or mission-planned separation events, all of which produce new objects in the vicinity of known objects.

The following sections provide a limited mathematical background in Finite Set Statistics and derivation of a simple CPHD spawning model, valid for the assumption that spawning is a Poisson process. A more thorough treatment of the subject is presented in Bryant et al. [8], which provides a detailed derivation based on point process theory and includes several variations of spawning processes. The work presented here is the direct output of a course taught by Dr. Daniel Clark, and with inputs from Daniel Bryant, Nicola Baresi, In-Kwan Park, Illán García Amor, and Nikhil Shetty, all of whom I thank for their contributions.

E.1 Mathematical Background

The derivation in this appendix makes use of Probability Generating Functionals (PGFs), which are related to the PHD and cardinality distribution through derivatives. The n -th order

functional derivative is taken with respect to a function in n directions, similar to the Gâteaux derivative. Given a PGFl $G[h]$, the PHD is found by taking the first derivative and setting $h = 1$.

$$\nu(\mathbf{x}) = \delta G(h; \delta_{\mathbf{x}})|_{h=1} \quad (\text{E.1})$$

To determine the cardinality from a PGFl, the n -th order Janossy density is first computed by taking the n -th order derivative and setting $h = 0$.

$$j^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) = \delta^n G(h; \delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})|_{h=0} \quad (\text{E.2})$$

Recalling the previous definition, Eq. (2.48), the cardinality of n targets is computed by integrating over the multitarget state space.

$$p(n) = \frac{1}{n!} \int j^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) d\mathbf{x}_1 \dots d\mathbf{x}_n \quad (\text{E.3})$$

In order to compute the derivatives above, the chain rule and product rule for functional derivatives need to be defined. The chain rule is based on Faá di Bruno's formula [13, 8].

$$\delta^n (F \circ G)(h; \eta_1, \dots, \eta_n) = \sum_{\pi \in \Pi(\eta_1, \dots, \eta_n)} \delta^{|\pi|} F(G(h); \delta^{|\omega|} G(h; \xi : \xi \in \omega) : \omega \in \pi) \quad (\text{E.4})$$

The product rule is based on Leibniz's formula [12, 8].

$$\begin{aligned} \delta^n (F(h)G(h); \eta_1, \dots, \eta_n) &= \sum_{k=0}^n \delta^{(n-k)} F(h; \pi) \delta^k G(h; \{\eta_1, \dots, \eta_n\} \setminus \pi) \\ \pi &\in \{\eta_1, \dots, \eta_n\} \\ |\pi| &= n - k \end{aligned} \quad (\text{E.5})$$

Given the above definition of Janossy density, any PGFl may be approximated with the following equation, even if its exact form is not known [13].

$$G(h) = \sum_{n=0}^{\infty} \frac{1}{n!} \int \prod_{i=1}^n h(\mathbf{x}_i) j^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) d\mathbf{x}_1 \dots d\mathbf{x}_n \quad (\text{E.6})$$

The n -th order derivative of this approximation is given by [13]

$$\delta^n G(h; \eta_1, \dots, \eta_n) = \sum_{m \geq n} \frac{1}{(m-n)!} \int \eta_1(\mathbf{x}_1) \dots \eta_n(\mathbf{x}_n) \prod_{i=n+1}^m h(\mathbf{x}_i) j^{(m)}(\{\mathbf{x}_1, \dots, \mathbf{x}_m\}) d\mathbf{x}_1 \dots d\mathbf{x}_m \quad (\text{E.7})$$

E.2 Derivation of Poisson Spawning Model

With the above definitions, it is possible to derive the cardinality prediction for the CPHD filter with spawning. The prediction of new GMM components to represent spawned objects is the same as for the PHD filter, provided in Vo et al. [80], and is therefore not included. Implementation details are discussed in the following section. This section provides the cardinality prediction for a Poisson spawning model, with the following additional assumptions:

1. Target birth is ignored.
2. Probability of survival is not state dependent, $p_S(\mathbf{x}) = p_S$.
3. Target survival is modeled as a Bernoulli process.

The predicted PGFl is assumed to take the form

$$G_{k|k-1}(h) = G_{k-1}(G_M(h|\cdot)G_\beta(h|\cdot)) \quad (\text{E.8})$$

and the target survival and spawning PGFls are defined by

$$G_M(h|\mathbf{x}) = 1 - p_S + p_S \int f_{k|k-1}(\mathbf{y}|\mathbf{x})h(\mathbf{y})d\mathbf{y} \quad (\text{E.9})$$

$$G_\beta(h|\mathbf{x}) = \exp \left[\beta \left(\int b(\mathbf{y}|\mathbf{x})h(\mathbf{y})d\mathbf{y} - 1 \right) \right] \quad (\text{E.10})$$

The density $f_{k|k-1}(\mathbf{y}|\mathbf{x})$ is the Markov transition density of individual targets, β is the Poisson mean rate for spawning, and spawned targets are distributed according to $b(\mathbf{y}|\mathbf{x})$. For simplicity, the substitution $G_c(h|\mathbf{x}) = G_M(h|\mathbf{x})G_\beta(h|\mathbf{x})$ is used at times.

To compute the Janossy density, find the n -th order derivative of $G_{k|k-1}(h)$ and set $h = 0$. The first step is to apply the chain rule, Eq. (E.4).

$$\delta^n G_{k|k-1}(h; \delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})|_{h=0} = \sum_{\pi \in \Pi(\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})} \delta^{|\pi|} G_{k-1} \left[G_c(0|\cdot); \delta^{|\omega|} G_c(0|\cdot; w) : w \in \pi \right] \quad (\text{E.11})$$

The exact form of the prior PGFl $G_{k-1}(\cdot)$ is not known, so it is approximated with Eq. (E.6). The derivative is then given by Eq. (E.7), with the following substitutions for the function h and

directions η_i .

$$h(\mathbf{x}_i) = G_c(0|\cdot) = e^{-\beta}(1 - p_S) \quad (\text{E.12})$$

$$\eta_i(\mathbf{x}_i) = \delta^{|w_i|} G_c(0|\mathbf{x}_i; w_i) \quad (\text{E.13})$$

The Janossy density is now given by

$$\begin{aligned} j_{k|k-1}^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) &= \sum_{\pi \in \Pi(\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})} \sum_{m \geq |\pi|} \frac{1}{(m - |\pi|)!} \\ &\times \int \delta^{|w_1|} G_c(0|\mathbf{x}_1; w_1) \dots \delta^{|w_{|\pi|}|} G_c(0|\mathbf{x}_{|\pi|}; w_{|\pi|}) \\ &\times \prod_{l=|\pi|+1}^m \left[e^{-\beta}(1 - p_S) \right] j_{k-1}^{(m)}(\{\mathbf{x}_1, \dots, \mathbf{x}_m\}) d\mathbf{x}_1 \dots d\mathbf{x}_m \end{aligned} \quad (\text{E.14})$$

Assuming the targets are independent and identically distributed (i.i.d.), the prior Janossy density

$j_{k-1}^{(m)}(\cdot)$ can be replaced by

$$j_{k-1}^{(m)}(\{\mathbf{x}_1, \dots, \mathbf{x}_m\}) = m! \cdot p_{k-1}(m) \prod_{j=1}^m s_{k-1}(\mathbf{x}_j) \quad (\text{E.15})$$

Substituting and reorganizing terms produces

$$\begin{aligned} j_{k|k-1}^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) &= \sum_{\pi \in \Pi(\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})} \sum_{m \geq |\pi|} \frac{m!}{(m - |\pi|)!} p_{k-1}(m) \\ &\times \int \prod_{i=1}^{|\pi|} \delta^{|w_i|} G_c(0|\mathbf{x}_i; w_i) \\ &\times \prod_{l=|\pi|+1}^m e^{-\beta}(1 - p_S) \prod_{j=1}^m s_{k-1}(\mathbf{x}_j) d\mathbf{x}_j \end{aligned} \quad (\text{E.16})$$

The final product has terms in the range $j \in [1, m]$ which may be split among the first two products, which cover the ranges $i \in [1, |\pi|]$ and $l \in [|\pi| + 1, m]$ respectively. Additionally, because \prod and \int are linear operators, their order may be switched. The equation can therefore be rewritten as

$$\begin{aligned} j_{k|k-1}^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) &= \sum_{\pi \in \Pi(\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})} \sum_{m \geq |\pi|} \frac{m!}{(m - |\pi|)!} p_{k-1}(m) \\ &\times \prod_{i=1}^{|\pi|} \left[\int \delta^{|w_i|} G_c(0|\mathbf{x}_i; w_i) s_{k-1}(\mathbf{x}_i) d\mathbf{x}_i \right] \\ &\times \prod_{l=|\pi|+1}^m \left[\int e^{-\beta}(1 - p_S) s_{k-1}(\mathbf{x}_l) d\mathbf{x}_l \right] \end{aligned} \quad (\text{E.17})$$

The derivative $\delta^{|w_i|}G_c(0|\cdot; w_i)$ is found by applying the product rule, Eq. (E.5). At this point, all derivatives will be given in terms of the n -th order and appropriate substitutions will be made later.

$$\begin{aligned} \delta^n(G_M G_\beta)(h|\cdot; w) &= \sum_{k=0}^n \delta^{(n-k)} G_M(h; \pi) \delta^k G_\beta(h; \{\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n}\} \setminus \pi) \\ \pi &\in \{\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n}\} \\ |\pi| &= n - k \end{aligned} \quad (\text{E.18})$$

Given the Bernoulli target survival process, only two terms from the summation are needed. This can be seen from the derivatives of Eq. (E.9).

$$\delta^{(n-k)} G_M(h; \pi) = \begin{cases} 1 - p_S & \text{if } n - k = 0 \\ p_S \cdot f_{k|k-1}(\mathbf{x}|\cdot) & \text{if } n - k = 1 \\ 0 & \text{if } n - k > 1 \end{cases} \quad (\text{E.19})$$

Therefore only the terms $k = n$ and $k = n - 1$ are retained. The derivative in Eq. (E.18) becomes

$$\delta^n(G_M G_\beta)(h|\cdot; w)|_{h=0} = \left[G_M(h|\cdot) \delta^n G_\beta(h; \delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n}) + \delta G_M(h; \pi) \delta^{(n-1)} G_\beta(h; \{\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n}\} \setminus \pi) \right]_{h=0} \quad (\text{E.20})$$

The second term must consider all subsets $\{\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n}\} \setminus \pi$ and since $|\pi| = 1$, can be replaced by the summation

$$\delta G_M(h; \pi) \delta^{(n-1)} G_\beta(h; \{\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n}\} \setminus \pi) = \sum_{i=1}^n \delta G_M(h; \delta_{\mathbf{x}_i}) \delta^{(n-1)} G_\beta(h; \{\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n}\} \setminus \delta_{\mathbf{x}_i}) \quad (\text{E.21})$$

The following individual derivatives are used to simplify Eqs. (E.20)-(E.21).

$$\begin{aligned} G_M(h|\cdot)|_{h=0} &= 1 - p_S \\ \delta^n G_\beta(h; \delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})|_{h=0} &= \beta^n e^{-\beta} \prod_{i=1}^n b(\mathbf{x}_i|\cdot) \\ \delta G_M(h; \delta_{\mathbf{x}_i})|_{h=0} &= p_S \cdot f_{k|k-1}(\mathbf{x}_i|\cdot) \\ \delta^{(n-1)} G_\beta(h; \{\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n}\} \setminus \delta_{\mathbf{x}_i})|_{h=0} &= \beta^{n-1} e^{-\beta} \prod_{\substack{j=1 \\ j \neq i}}^n b(\mathbf{x}_j|\cdot) \end{aligned}$$

The final result of applying the product rule is the n -th order derivative

$$\delta^n(G_M G_\beta)(h|\cdot; w)|_{h=0} = \beta^n e^{-\beta} (1 - p_S) \prod_{i=1}^n b(\mathbf{x}_i|\cdot) + \beta^{n-1} e^{-\beta} p_S \sum_{i=1}^n f_{k|k-1}(\mathbf{x}_i|\cdot) \prod_{\substack{j=1 \\ j \neq i}}^n b(\mathbf{x}_j|\cdot) \quad (\text{E.22})$$

Returning to the Janossy density Eq. (E.17), and again switching the order of \prod and \int where appropriate produces

$$\begin{aligned} j_{k|k-1}^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) &= \sum_{\pi \in \Pi(\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})} \sum_{m \geq |\pi|} \frac{m!}{(m - |\pi|)!} p_{k-1}(m) \\ &\times \prod_{i=1}^{|\pi|} \left[\beta^{|\mathbf{w}_i|} e^{-\beta} (1 - p_S) \prod_{j=1}^{|\mathbf{w}_i|} \int b(\mathbf{x}_j|\mathbf{x}_i) s_{k-1}(\mathbf{x}_i) d\mathbf{x}_i \right. \\ &+ \beta^{|\mathbf{w}_i|-1} e^{-\beta} p_S \sum_{j=1}^{|\mathbf{w}_i|} \int f_{k|k-1}(\mathbf{x}_j|\mathbf{x}_i) \prod_{\substack{l=1 \\ l \neq j}}^{|\mathbf{w}_i|} b(\mathbf{x}_l|\mathbf{x}_i) s_{k-1}(\mathbf{x}_i) d\mathbf{x}_i \left. \right] \\ &\times \prod_{l=|\pi|+1}^m \left[e^{-\beta} (1 - p_S) \int s_{k-1}(\mathbf{x}_l) d\mathbf{x}_l \right] \end{aligned} \quad (\text{E.23})$$

Assume that all the distributions $f_{k|k-1}(\cdot)$, $b(\cdot)$, and $s_{k-1}(\cdot)$ are Gaussian. From Lemma 1 in Reference [80], the integral of a product of Gaussian distributions is another Gaussian distribution. The following simplifications can therefore be made.

$$\int b(\mathbf{x}_j|\mathbf{x}_i) s_{k-1}(\mathbf{x}_i) d\mathbf{x}_i = p_g(\mathbf{x}_j) \quad (\text{E.24})$$

$$\int f_{k|k-1}(\mathbf{x}_j|\mathbf{x}_i) \prod_{\substack{l=1 \\ l \neq j}}^{|\mathbf{w}_i|} b(\mathbf{x}_l|\mathbf{x}_i) s_{k-1}(\mathbf{x}_i) d\mathbf{x}_i = \prod_{\substack{l=1 \\ l \neq j}}^{|\mathbf{w}_i|} p_g(\mathbf{x}_j) \quad (\text{E.25})$$

$$\int s_{k-1}(\mathbf{x}_l) d\mathbf{x}_l = 1 \quad (\text{E.26})$$

The final Janossy density is therefore

$$\begin{aligned} j_{k|k-1}^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) &= \sum_{\pi \in \Pi(\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})} \sum_{m \geq |\pi|} \frac{m!}{(m - |\pi|)!} p_{k-1}(m) \\ &\times \prod_{i=1}^{|\pi|} \left[\beta^{|\mathbf{w}_i|} e^{-\beta} (1 - p_S) \prod_{j=1}^{|\mathbf{w}_i|} p_g(\mathbf{x}_j) + \beta^{|\mathbf{w}_i|-1} e^{-\beta} p_S \sum_{j=1}^{|\mathbf{w}_i|} \prod_{\substack{l=1 \\ l \neq j}}^{|\mathbf{w}_i|} p_g(\mathbf{x}_j) \right] \\ &\times \prod_{l=|\pi|+1}^m \left[e^{-\beta} (1 - p_S) \right] \end{aligned} \quad (\text{E.27})$$

The cardinality is computed from Eq. (E.3), which involves integrating over the multitarget state space. The only terms with dependency on the state are the Gaussian distributions, which are integrated out as follows:

$$\prod_{j=1}^{|w_i|} \int p_g(\mathbf{x}_j) d\mathbf{x}_j = \prod_{j=1}^{|w_i|} 1 = 1 \quad (\text{E.28})$$

$$\sum_{j=1}^{|w_i|} \prod_{\substack{l=1 \\ l \neq j}}^{|w_i|} \int p_g(\mathbf{x}_j) d\mathbf{x}_j = \sum_{j=1}^{|w_i|} 1 = |w_i| \quad (\text{E.29})$$

The predicted cardinality is therefore

$$\begin{aligned} p_{k|k-1}(n) &= \frac{1}{n!} \int j_{k|k-1}^{(n)}(\{\mathbf{x}_1, \dots, \mathbf{x}_n\}) d\mathbf{x}_1 \dots d\mathbf{x}_n \\ p_{k|k-1}(n) &= \sum_{\pi \in \Pi(\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_n})} \sum_{m \geq |\pi|} \frac{m!}{n!(m - |\pi|)!} p_{k-1}(m) \\ &\times \prod_{i=1}^{|\pi|} \left[\underbrace{\beta^{|w_i|} e^{-\beta} (1 - p_S)}_{\text{Spawn and Die}} + \underbrace{|w_i| \beta^{|w_i|-1} e^{-\beta} p_S}_{\text{Spawn and Survive}} \right] \\ &\times \prod_{l=|\pi|+1}^m \left[\underbrace{e^{-\beta} (1 - p_S)}_{\text{Target Death}} \right] \quad : w_i \in \pi \end{aligned} \quad (\text{E.30})$$

E.3 Implementation and Results

To demonstrate the approach, a simulation is included, in which one object is tracked and two new objects are added. The simulation follows the same basic structure as Ch. 2, using dense measurements over a 12 hour window, and with objects randomly initialized near the EchoStar1 satellite. At the beginning of the observation window, one object is present, and two spawned objects are added at the two hour mark. The filter is assumed to have perfect knowledge of the time of the spawning event and the filter parameters are the same as used in Ch. 2, with the exception that there are no missed detections or clutter measurements. While this is an overly idealized case for SSA, it demonstrates that the spawning model functions correctly.

The cardinality prediction makes use of Eq. (E.30) with a mean rate of $\beta = 2$ spawned targets and no target birth. With this prediction for cardinality, the PHD prediction need only add one GMM component for each new target, near the original target and offset by some deviation in

the state vector. In this simplified scenario, the deviation in state vector is assumed to be known exactly, which is unrealistic in general, but again serves to demonstrate the functionality of the approach.

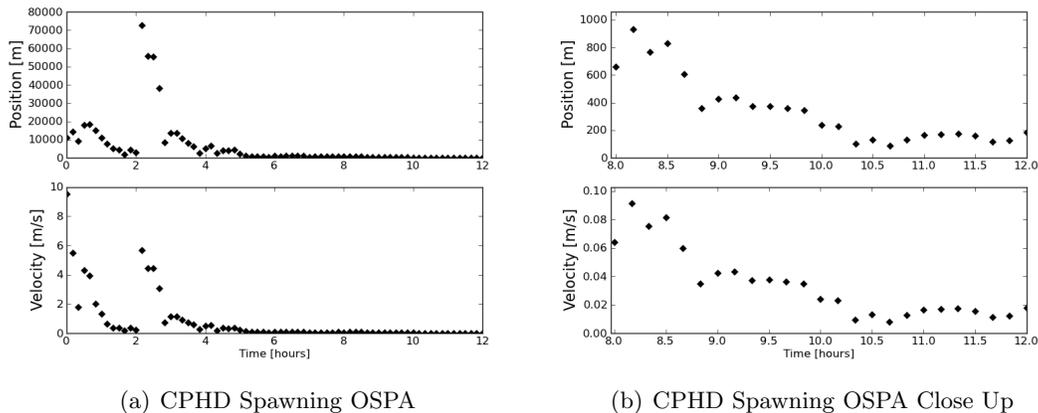


Figure E.1: CPHD Spawning OSPA Results

Figure E.1 provides the OSPA results for the spawning test case. At the two hour mark, there is an increase in the state errors due to the new objects, resulting in a position error of 70 km. By the end of the observation window, the filter has converged to estimation errors of around 200 m, which is similar to that observed in the Ch. 2 simulations.

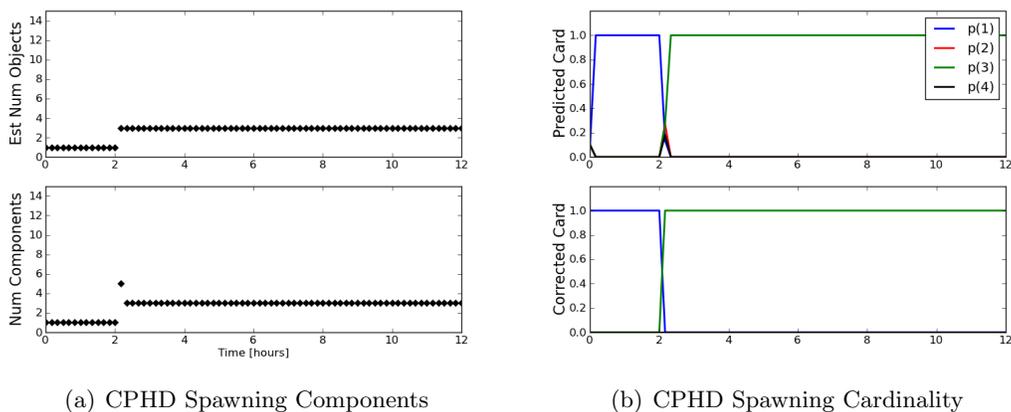


Figure E.2: CPHD Spawning Cardinality Results

Figure E.2 provides results related to the estimated number of objects and cardinality. The

filter is quickly able to identify the correct number of objects and converges to using a single GMM component to represent each object. The extraneous components at the time of target spawning are due to the AEGIS implementation, in which the propagation step is split into smaller time intervals and spawning GMM components are generated in each of the shorter intervals, instead of only once at the end. This is an implementation choice, and could be changed in the future, however, the filter quickly corrects itself and converges to three components. In longer data gaps, if the time of spawning is not known, it may be best to assume spawning events at each of the intermediate times and allow the filter to update when measurements are received, as is the case here. The cardinality prediction and correction function as desired, immediately identifying three objects after the spawning event.

While overly simplified, this simulation demonstrates that the cardinality prediction equation derived in Eq. (E.30) can be used to model new targets in the CPHD filter. Extension of this method to more realistic spawning models has already been undertaken [8], and offers a promising direction for future research. The approach may provide advantages over IOD techniques by using knowledge of existing object states to quickly and accurately initiate tracks for newly generated objects in orbit.