Graduate School
UNIVERSITY OF COLORADO **BOULDER**

Thesis Approval Form

Student Name Sam Wishnek Student ID *********

Doctoral Degree in Aerospace Engineering Sciences

Thesis Title:

Optimal Information Theoretic Techniques for Electro-Optical Space Domain Awareness

Have you performed research involving human subjects which requires approval from the Institutional Review Board (IRB)? ___ Yes _X_ No

IRB Protocol Number _____

Have you used live animals, animal tissue, or observational animal work which requires approval from the Institutional Animal Care and Use Committee (IACUC)? ___ Yes _X_ No

IACUC Protocol Number _____

Attach the final copy of thesis/dissertation for committee review. While formatting changes may be requested by the Graduate School, the content of the attached document should be final.

Approvals:

Committee Chair Name Marcus Holzinger

Signature DocuSigned by: Marcus Holzinger 5038EBEB28BA425... Date Signed 11/15/2021

Committee Member Name Nisar Ahmed

Signature DocuSigned by: Nisar Ahmed 14B89CCD093F471... Date Signed 11/15/2021

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.

Optimal Information Theoretic Techniques for Electro-Optical Space Domain Awareness

by

Samuel W. Wishnek

B.A., University of California, Berkeley, 2016

M.S., University of Colorado at Boulder, 2020

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

2021

This thesis entitled:
Optimal Information Theoretic Techniques for Electro-Optical Space Domain Awareness
written by Samuel W. Wishnek
has been approved for the Department of Aerospace Engineering Sciences

_____

Prof. Marcus J. Holzinger

_____

Prof. Nisar Ahmed

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.

Wishnek, Samuel W. (Ph.D., Aerospace Engineering)

Optimal Information Theoretic Techniques for Electro-Optical Space Domain Awareness

Thesis directed by Prof. Marcus J. Holzinger

With high and growing public interest from key stake holders in exploring and using the space domain. The tools and techniques of space domain awareness (SDA) have seen growing interest and development. The field of space domain awareness covers the methods for detecting, tracking, and characterizing the states of space objects. These tools include every step in the process for building up estimates of these states and are essential for mission planning and actualization as well as collision avoidance. This starts with the tools for data acquisition and continues on through orbit determination and state estimation. While the field of SDA has seen incredible growth in recent years, there remains significant room for improvement.

Information optimal approaches extend beyond just providing results. They take into the account the volume of information in the ingested data and provide estimates and the corresponding uncertainty. This allows information optimal approaches to build a result that uses all the available information in the data without communicating a false level of certainty in the resulting estimate.

This research discusses new methods and algorithms for information-optimal space domain awareness. The presented methods extend from data collection to state estimation. While some of the methods are generally applicable to many different scenarios, there is a focus on electro-optical data. These methods enable more and robust methods for understanding the space domain both in and beyond the near-Earth environment.

Acknowledgements

My advisor, Professor Marcus Holzinger was a tremendous help in this research. His support was invaluable in developing this research and discovering new avenues for approaching the various problems and roadblocks we encountered along the way.

For much of my time at CU Boulder, my work was funded under a contract through Ball Aerospace. Through this contract I worked with Patrick Handley, Sue Hagerty, and Jacob Griesbach. Their advice and direction was instrumental in building up the latter half of this body of research including both of the AMOS presentations.

My peers, Shez Virani, Daniel Aguilar Marsillach, and Sam Fedeler were fantastic to work with at both Georgia Tech and CU Boulder. Their assistance was essential for collecting the data used to observationally verify this research.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Problems in the Observation to Estimation Pipeline

1.1 Space Domain Awareness

Over the past few decades, the near-Earth environment has become increasingly important for supporting modern civilization through the novel application of technology to enhance communication, science, and security among many other interests. Accordingly, over the years the environment has seen greater use with more government and non-government entities placing satellites into orbit. This has led to an increased demand for accurate and up-to-date knowledge of the states of objects in the space environment [2]. This is both for tracking the state of operational spacecraft to inform future maneuvers as well as the states of other space objects to prevent collisions that could disrupt missions.

Space domain awareness (SDA) is the identification, tracking, and orbit characterization of space objects [3]. There are a wide variety of tools and techniques in SDA. There are radio and optical frequency sensors, and these sensors can be based on ground stations or other spacecraft [4]. For the variety of different sensor types, there is an even broader spectrum of available computational tools for processing the resulting data. These are the filters and estimators that can extract information regarding a spacecraft's past, present, or likely future states. The techniques vary in their effectiveness on different criteria. Collectively, the various algorithms form a repertoire of tools that can be used to perform SDA operations. The optimal tools in a given scenario depend on the available information and operation goals.

There are a variety of SDA algorithms for the corresponding variety of feasible scenarios.

These tools span the full process of state estimation from extracting quantitative measurements from raw data to combining these measurements to build initial characterizations of the target orbit to narrowing in on the precise state estimate that can be propagated to find predictions of past and future states. With no a-priori knowledge of a target object's state, a single measurement of a space object with optical sensors is insufficient to build a full state. These sensors tend to provide two or four element measurements while a full translational state has six degrees of freedom. In general, angles-only observations provide two element measurements while angle and angle rate observations provide four element measurements. In either case, several measurements would be required in order to build a full state estimate. When this minimum is barely met, the typical approach is to apply an initial orbit determination (IOD) algorithm to estimate the state [5]. IOD algorithms often rely on basic assumptions about the target object's orbit and measurements such as pure Keplerian motion or short time of flight between observations. Accordingly, IOD algorithms tend to be less accurate than the alternative. Precise orbit determination (POD) algorithms are used when there are more measurements to include in the estimate, and the state is significantly overdetermined. Linearized, extended, and unscented Kalman filters fall into this category as well as particle filters. POD algorithms can use the greater volume of available data to provide more accurate estimates of the state. Some of these algorithms can also compensate for unmodeled dynamics as more data is ingested into the algorithm.

Space domain awareness is an evolving field and there are a number of ways the existing toolset can be improved and augmented with new approaches. With some sensors, the existing methods for extracting quantitative measurements are not information-optimal in that the derived measurements may not accurately reflect all available data in the observation and the measurement covariance matrix may not accurately reflect the true uncertainty in the derived values. Further down the process line, there is room for improvement in the available IOD methods to better apply the information available in the collected data.

Thesis Statement: Electro-optical space domain awareness tools can benefit from a novel set of optimal information theoretic techniques that can be used to extract state estimates

from measurements while simultaneously handling ambiguity and rigorously quantize confidence measures.

The contributions to the state of the art break down into four major categories. Each of these corresponds to a step in the process of building a full state estimate from initial data collection to orbit determination. The motivation and impact of the these topics are broken down in the next few sections. Broadly speaking, each contribution helps add rigor and precision to the full process of orbit determination. These advancements may individually provide small improvements to the accuracy of estimates and reliability of methods. However, the described approaches can be used together to enable orbit determination in broader realms with tools that bring powerful capabilities to stake holders in the commercial, scientific, and military domains. This helps close in on a more complete set of tools to reduce the set of infeasible problems in space domain awareness.

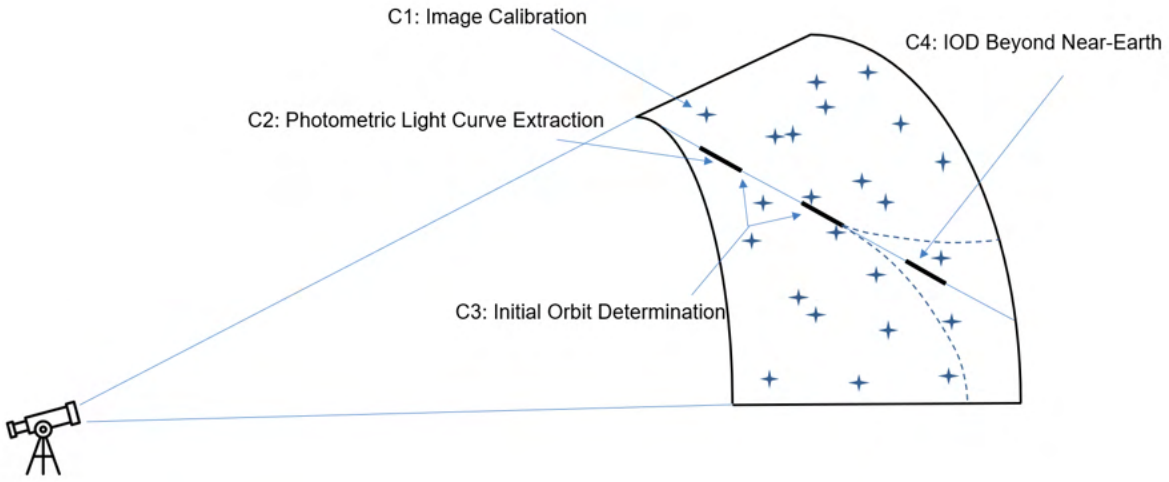

Figure 1.1: Diagram showing the four major contributions of this research and how they relate to each other in optical data.

1.2 The Complications of Ultra-Wide Angle Lenses

For electro-optical data collection, there are a variety of tools available that each specialize in accomplishing different tasks and enabling different objectives. Narrow field-of-view telescopes can

provide highly accurate relative state measurements of target objects with minimal uncertainty. Beyond this, some can be used for resolved imaging of space objects and this can support tasks such as attitude determination and predictive modeling for non-gravitational forces on the target object. However, these narrow field-of-view instruments can be limited by the same aspect that gives them these capabilities. The narrow field-of-view limits their ability for a single instrument to characterize and track multiple objects. This is particularly true when multiple objects of interest are observable at the same time but not nearby to one another. These instruments also struggle at uninformed detection tasks as they can only observe a small portion of the night sky at a time.

An alternative option more capable with these tasks are wide angle and particularly, ultra-wide field-of-view telescopes. These are instruments with fields-of-view over sixty degrees [6]. While some of these may fall behind in their ability to provide high accuracy state measurements and they are largely restricted to unresolved data collection, they are highly capable at the tasks narrow-field of view instruments fall short in. Their fields-of-view can capture a large portion of the night sky and even fixed bore-sight cameras can be reliably used for SDA data collection. They excel at uninformed detection and tracking tasks as well as multi-target tracking as it is much more likely an observable target object will enter their view cone than for a narrow field-of-view counterpart [7].

The benefits of wide and ultra-wide field-of-view instruments have already attracted interest from a variety of sources. One of the most notable instruments is the Vera C. Rubin Observatory. This observatory, also known as the Large Synoptic Survey Telescope (LSST), will gather several types of data products for a variety of tasks including all-sky surveys, studies on survey cadence, and transient detection [8]. Scheduled for first light in 2022, this observatory capability to survey the whole night sky separate it from contemporary narrow field-of-view telescopes like the Giant Magellan Telescope, Extremely Large Telescope, and Thirty Meter Telescope. There are also other smaller telescope systems intended for sky surveys such as the Dragonfly Telephoto Array [9].

One of the major issues faced by observational systems with ultra-wide fields-of-view is distortion across the image plane induced by the optical system [10]. The light the sensor receives

must be significantly bent to reach the imaging surface. Distortion can arrive as a simple consequence of the required manipulation of the light to reach the sensor or it can be induced by small imperfections in the optical system that lead to deviations from the expected distortion of the optical system. These issues are some of the major barriers to the wider adoption of ultra-wide angle sensors that would otherwise offer a relatively low-cost solution for passively monitoring a large portion of the night sky [11].

Existing algorithms for image calibration to remove this distortion involve fitting the observed images to a model and estimating the coefficients that define the distortion from a series of calibration images. Popular methods for this include botht the Brown and division models for image calibration [12] [13]. These approaches work best for cases with less severe distortion. However, there is an opportunity to improve over these methods for the case of ultra-wide angle lenses that observe the night sky. This is empirical calibration. Rather than fitting the data to a model to estimate coefficients, stars can be used as reference points to gather information about the camera's distortion and then build a model by interpolating between points with known distortion.

This approach to image calibration offers several advantages over the function-fitting algorithms. First, it does not require manual intervention to build the calibration map as it uses the same type of data it collects to observe space objects as it uses to calibrate itself. Second, the empirical approach can account for complex and irregular distortion that may violate the underlying symmetry assumptions of a function-based model. This allows it to better model irregular errors from manufacturing defects. Lastly, this approach can allow a calibration to be automatically updated or reacquired as more data is collected by the sensor.

This contribution enables the broader and more reliable use of ultra-wide field-of-view sensors. These may have been previously considered too inaccurate with data products too difficult to extract useful measurements from [11]. These kinds of lenses and sensors are available to the broader public at a much lower cost than astronomy-focused alternatives. Implementation of this approach can open up all-sky data collection to more budget constrained observers as well as making systems with more sensors feasible.

Contribution 1: An algorithm for empirically calibrating night sky images taken with ultra-wide angle lenses in order to accurately report the observed relative position of imaged space objects.

1.3 Light Curve Measurement and Uncertainty

When an electro-optical sensor images a non-tracked moving object, that object shows up as a streak in the resulting image. The light from the source slides across the observing sensor and the effect is integrated over the time the sensor is active. For a fixed-bore-sight camera it will almost always be the case that an observed space object will show up as a streak in the image. The length of this streak depends on the integration time and relative angular speed of the object. If the target objects grows dimmer or brighter while it streaks across the frame, this can show up as a dimming or brightening of the streak. Accordingly, the streak encodes information about the target object's photometry. That can provide additional insight into aspects of the target's attitude or relative position with respect to the observer, Sun, and Earth.

Pulling out the photometric information encoded in a streak produces a photometric light curve. The observed changes in brightness as an object streaks across an imaging frame can reveal other aspects of the target object's state. For spacecraft, some parts will have different albedos and reflective cross sections. As a satellite rotates with respect to the observer, the object may reflect more or less light. For example, solar panels tend to be highly reflective [14]. This will then be reflected in the observed light curve. In these cases, the photometric light curve can reveal information about the target object's attitude [15]. Alternatively, as the spacecraft passes into the Earth's penumbra, the Sun will be occluded and there will be less incident light on its surface. The observed reflected light will fall off accordingly and this can be indicated in the observed light curve.

The usual approach for estimating the light curve based on an observed streak is relatively simple. The streak is located in the image and broken down into an arbitrary number of segments. The pixels within each segment are added together to integrate the light to find the intensity of

each segment. These segments are then plotted as a piecewise light curve to model the photometry [16] [17].

This approach leaves several impactful choices as arbitrary guesses and lacks a meaningful way to ensure the estimated streak optimally matches the observation. Assuming that the observed object has any point spread function on the sensor, this approach fails to account for the geometry of how the streak was produced. The point spread function slides across the sensor as the image integrates the light. This means that later portions of the streak do represent later times, but there is some overlap as the point spread function from a moment before and a moment after overlap on the same pixel. Next there is the question of how to break down the streak into segments. If too small a number of segments is chosen then there is usable information contained in the streak that is not included in the estimate. On the other hand, if the number of segments is too large, the resulting light curve will have values for unobservably small segments of time and indicate a false level of certainty in the estimated light curve. Finally, this approach offers no method for describing the uncertainty in the estimate. This makes it difficult to rigorously assess the conclusions derived by a resulting estimate.

In this contribution, a rigorous, information theoretic approach is offered for estimating the photometric light curve from an observed streak. The approach develops a method for breaking down the streak into an optimal number of segments to maximize the information used in the result. This is derived from a more complete model of how a streak is generated as the point spread function slides across the imaging sensor. This approach also develops a method for finding the covariance of the resulting estimate so the level of certainty in the results can be properly understood. In addition, a simplified analytical approach for finding the optimal number of streak segments is derived for cases that meet certain geometric assumptions.

This approach offers a reliable and robust method for estimating the photometric light curve of observed streaking space objects. The developed method can be applied to extract all the available information in the streak so the estimate can most closely reflect the available data. This provides a clear and objective way to find photometric light curves from streaks and removes ad

hoc and arbitrary choices in finding these values.

Contribution 2: A pair of methods for determining the information-optimal way to discretize an observed space object streak into segments in order to fully extract the available photometric information in the streak and find the covariance of the time resolved light curve estimate.

1.4 Initial Orbit Determination from Streaks

Initial orbit determination is a method by which a minimal set of observations can be used to generate a full dynamical state for an orbiting space object. Within the realm of initial orbit determination from optical data, angles-only methods generally use three observations while methods with angles and angle rates use two. That is because the former includes two element measurement vectors so there are six values that go into finding the full six-element dynamical state. The latter uses four element states and a minimum of two observations are required for the full dynamical state to be observable [18]. There are a variety of algorithms available for initial orbit determination. Each has strengths and weaknesses depending on the specific use case and orbit determination goals.

There are many different initial orbit determination algorithms. Most of these algorithms intended for optical data work with angles-only data. For example, the popular Laplace, Gauss, Gooding, and double-r algorithms are all intended for angles-only data. Even within angles-only cases these methods differ in their capabilities and implementation requirements. Methods such as Laplace and Gauss tend to struggle with targets in low-Earth orbit as these methods perform best when the angle between observations is small. This narrows the feasible time windows between observations that allow these methods to remain useful [19]. Other angles-only methods like P.R. Escobal's double-r method work well with large angles between measurements. However, the iterative double-r algorithm has a relatively tight constraint on the initial guess for the target ranges that allows the method to converge to the correct solution [20]. The popular Gooding method often fails for space-based observers [19][21]. All these methods generally assume that the central body acts as a gravitational point source and no other forces are influencing the dynamics. For target

objects in the near-Earth environment, this is generally a safe assumption for finding a reasonably accurate initial estimate. The results can then be fed into a precise orbit determination algorithm to refine the estimate and generate a state covariance matrix.

Since most methods for precise orbit determination require an initial guess to start from, initial orbit determination methods are an essential step in finding a full state estimate. They can take measurements and generate a close initial guess that can then be used to seed the precise orbit determination algorithm. More universal and robust approaches to initial orbit determination can broaden the scope of problems with feasible solutions.

Since the various methods work better in different domains, an operator must typically have some knowledge of the target in advance or implement multiple methods to make sure the most relevant method is applied. Even then, some the algorithms fail by providing inaccurate results so it may not be clear which method worked and which failed. There are also very few methods that accept measurements with both angles and angle rates. If the intention is to build a state estimate from two observations with angles and angle rates, some domains that are feasible for angles-only data are infeasible due to a lack of algorithms that excel in that domain.

The proposed solution to this problem is an optimization-based approach to initial orbit determination. The method formulates how well a solution fits the measurements as a cost function and adjusts the unmeasured state components to minimize the error from the observations. It accepts input measurements of both angles and angle rates to naturally work with measurements derived from streaks. The approach also applies admissible regions as a penalty function in order to restrict the search space and more efficiently converge to a solution. The algorithm can also solve for the set of feasible solutions when there are multiple states that fit the observations such as in the case of multiple orbits.

This contribution provides a robust method for initial orbit determination that works under a wide variety of circumstances. This includes both orbiting and ground-based observers as well as cases where the target object has an orbit closer to the Earth than the observer. This is a case that failed for many other algorithms. The approach successfully converges for both long and short

times of flight as well.

Contribution 3: A novel IOD method that accepts both angles and angle-rates to provide a state estimate that remains accurate for a broad range of orbital regimes as well as long delays between observations without an initial guess.

1.5 Initial Orbit Determination in Cislunar and Beyond

In the near-Earth environment, the dynamics of an orbiting object are dominated by the gravitational influence of the Earth. The gravitational pull of the Earth is powerful enough that modeling the dynamics as a simple two-body point source problem is sufficient for finding a near estimate for the purpose of initial orbit determination. Accordingly, most existing methods for initial orbit determination rely on the two-body assumption to efficiently find a state estimate [18]. Beyond the near-Earth environment, this assumption breaks down. The gravitational influence of the Moon, other solar system bodies, and non-gravitational influences become more significant to the dynamics as the dominance of the gravitational influence of the Earth falls off.

Initial orbit determination is largely restricted to cases well-approximated by the two-body assumption. Under these simplified dynamics, there are a number of different approaches to analytically or computationally finding the full six-element dynamical state. The most common IOD algorithms such as Laplace, Gauss, Gooding, and double-r are all derived from the assumptions of two-body dynamics [18]. Under these assumptions, these algorithms can very efficiently converge to close estimates of the full dynamical state. However, these methods will not work if the two-body assumption breaks down as the assumptions they are based on no longer hold.

An approach for performing initial orbit determination outside of the near-Earth domain would enable stake holders for space beyond the near-Earth environment a better understanding of the environment and the objects within it. With NASA looking to revisit the Moon in the coming years as well as new interest from commercial and military sectors in the cislunar domain, new and improved approaches for better understanding and managing the environment could enable many future missions [22]. With IOD an essential step in the process of space object characterization

and tracking, an algorithm that extends this capability beyond the near-Earth environment could significantly improve SDA capabilities in this new domain.

To solve this problem, an algorithm is developed that uses an optimization-based approach similar to the previous contribution, but works with a differential equation integrator to estimate the state subject to arbitrary dynamics. This is implemented with a high-fidelity orbit propagator that includes a variety of forces on the target object so the approach will converge to an accurate solution both inside and outside of the near-Earth environment. The cost function is based on the Mahalanobis distance, so the resulting cost is a meaningful measure of how accurately the state fits the observations it was derived from. This approach allows for the inclusion of process noise in building an estimate so as not to exclude potential solutions due to over confidence in the accuracy of the modeled dynamics. The approach comes with analytical necessary condition for convergence that allows some potential solutions to be excluded as infeasible.

Contribution 4: A method for performing initial orbit determination under arbitrary dynamics and enabling orbit determination in the cislunar domain and beyond.

1.6 Optical Observations with OmniSSA

The VADeR lab at the University of Colorado at Boulder has constructed a multi-sensor ultra-wide field-of-view optical system called the Omnidirectional Space Situational Awareness (OmniSSA) system. The goal of this hardware project is to approach the problem of uninformed object detection and state estimation using four ultra-wide angle sensors. These sensors are composed of ZWO ASI1600MM cameras with Rokinon FE14M-E lenses in order to achieve 115 degree field of view images of the night sky. The fields of view of the four sensor overlap entirely. The output images can be used together in order to improve detection signal-to-noise ratio limits [23]. The system is designed for long term autonomous operation and remote access in order to passively observe the night sky and perform state estimation on a large set of observable space objects.

OmniSSA has been shown to be able to detect objects as dim as magnitude 12 which allows it to observe geostationary and geosynchronous space objects. The system is zenith pointing, so

fast moving space objects such as those in low Earth orbit appear as streaks in the sensor frame for longer integration times.

OmniSSA is used to provide real data to observationally verify the algorithms applied in every step of this research. First, the space object streaks created by low Earth objects in OmniSSA's data provide a target for photometric light curve extraction. Next, the optical lens itself induces distortion in the image relative to the expected projection of the sky. This provides a distorted target to apply the ultra-wide angle lens calibration. The large datasets from leaving OmniSSA to run for several hours while observing a significant portion of the sky provide ample targets for testing the initial orbit determination approaches discussed in contribution three.

1.7 Contributions, References, and Papers

Reference	Calibration	Photometry	IOD	General IOD
Brown [12]	✓			
Fitzgibbon [13]	✓			
Shah et al. [24]	✓			
Wang et al. [25]	✓			
Stein [26]	✓			
Ai et al. [27]	✓			
Virtanen et al. [28]		✓		
Silha et al. [16]		✓		
Chote et al. [17]		✓		
Kurosaki et al. [29]		✓		
Gooding [21]			✓	
Escobar [20]			✓	
Karimi et al. [30]			✓	
Karimi et al. [31]			✓	
DeMars et al. [32]			✓	
Taff et al. [33]			✓	
Taff [34]			✓	
Gauss [35]			✓	
Milani et al. [36]				✓
Tommei et al. [37]				✓
Marruskin et al. [38]				✓
Farnocchia et al. [39]				✓
Holzinger [40]				✓
My Publication	Calibration	Photometry	IOD	General IOD
Conference [41]		SciTech 2020		
Conference [42]			AMOS 2020	
Conference [43]	AIAA 2021			
Conference [44]				AMOS 2021
Journal [45]			JAS	
Journal [46]	Submitted ASR			
Journal [47]		Submitted ASR		
Journal				Ready ASR

Chapter 2

Ultra-Wide Angle Lens Calibration

2.1 Image Calibration

Generally defined as lenses with field-of-view over sixty degrees, ultra-wide field-of-view optical sensors offer a number of advantages over their narrow field-of view counterparts in SDA [6]. Primarily, they allow fewer individual sensors to simultaneously monitor a large portion of the sky. So long as the target objects are bright enough to be detected by the sensor, a single ultra-wide field-of-view sensor can passively monitor a much larger portion of the sky than a narrow field-of-view alternative. This makes these sensors particularly useful for uninformed object detection and fixed zenith pointing applications [7].

One of the major issues for any optical system relying on ultra-wide angle sensors is the distortion induced by the optical system of large field-of-view sensors [10]. In particular, the standard approach for removing distortion in images is often not sufficient for accurately determining the inertial coordinates of an imaged object. Rather, these methods tend to be unable to account for all of the distortion in the image and the estimated inertial angular position of an imaged object tend to be less accurate accordingly. In-house testing has shown that many off-the-shelf methods for image calibration for these types of sensors either do not well approximate the observed distortion or fail to produce a solution at all.

Existing methods for image calibration tend to be functional in that the distortion is modeled as an analytical function of position in the image where the coefficients of the various terms in the function determine the overall distortion of the image [25]. For example, popular methods for

removing distortion from images such as the Brown model and division model are functional [12] [13]. There are also methods specific to certain types of ultra-wide angle lenses such as fish-eye lenses [24]. These methods often work well for lenses with less severe and more symmetric distortion. However, the ultra-wide angle lenses in some low-cost systems can have fairly severe distortion that may be asymmetric or irregular in a way the precludes these functions from accurately reproducing the distortion for any choice of coefficients [27] [48].

The proposed solution for this problem is to use an empirical approach to mapping the distortion across the imaged frame. The empirical approach is based on gathering a collection of points with known distortion across the frame and interpolating between them. This method can account for any geometry with any distribution of points and improves as more calibration points are included. The observed to catalog star assignment problem is an example of the generalized assignment problem and is NP-hard [49]. This research does not address the assignment problem itself, but implements existing solutions to study the impact of empirical solutions on optical calibration.

Specifically, this approach entails using the locations of stars in data images as reference points for the calibration. As this method uses the observed and catalog coordinates of stars in night-sky images, this empirical approach does not require the manual collection of specific calibration images. Rather, the data used to calibrate the imaging system is the same type of data that is used to detect and characterize the orbits of near-Earth objects. This both reduces the amount of manual setup required for an optical system as well as allows calibration maps to be easily regenerated in the event of a suspected or known change to the optical system that might impact the accuracy of an existing map. For example, a lens replacement or focus adjust. The proposed methods leverage the Delaunay triangulation as well as the Savitzky-Golay filter in order to build the calibration map under different regimes [50] [51].

The proposed methods for ultra-wide field-of-sensor calibration offer several advantages over traditional methods. These empirical approaches can can be applied without manual intervention for sky-pointing cameras and can compensate for irregular or asymmetric distortions in the optical

system. The two proposed methods function better in different scenarios, and each is compared in their performance against data collected from an observation system. In addition, the Delaunay approach, which does not work as well for the given scenario, is assessed with simulated data for a case that better fits its capabilities.

2.2 Theory

2.2.1 State of the Art

Popular methods for image calibration are typically defined by vector-valued functions that take the coordinates of a pixel and output the calibrated coordinates. The function depends on the position of a point with respect to the image frame as well as a set of coefficients. These coefficients are determined by taking calibration images of a pre-determined target such as a checkerboard pattern with a known size and performing a least squares regression to find the set of coefficients that most closely match the model with the observed distortion. For example, equations (2.1 and (2.2 show the image frame horizontal and vertical coordinate calibration based on the Brown-Conrady distortion model [12]. In this case, the variables labeled as P_i where i is a positive integer are the coefficients found via the least squares regression. These are the terms that depend on the specific optical system. The remaining terms, x, y, and r are the horizontal, vertical, and radial positions of a given point in the image frame respectively with respect to the image center. Note that in these equations r is the magnitude of the (x, y) vector.

$$\Delta x = \left[P_1(r^2 + 2x^2) + 2P_2xy\right]\left[1 + P_3r^2 + P_4r^4 + ...\right] \tag{2.1}$$

$$\Delta y = \left[2P_1xy + P_2(r^2 + 2y^2)\right]\left[1 + P_3r^2 + P_4r^4 + ...\right] \tag{2.2}$$

These equations intrinsically assume a symmetry in the optical system about a boresight. Should the sign of the x and y values flip, the result will be unchanged. This assumption, while broadly near accurate for many optical systems, is not necessarily true. As an optical system is a practical application of real objects, there can be asymmetries that a low-order functional expansion

approach cannot account for.

Similarly, the division model developed by A.W. Fitzgibbon is shown in equations (2.3) and (2.4) [13]. This form solves for the undistorted coordinates rather than the distortions. These undistorted coordinates are labeled (x_{real}, y_{real}). This model also has intrinsic symmetry about the central point. In both cases, it is trivial to shift the location of this central point by modifying what coordinates the input values are relative to. However, the symmetry cannot be removed, and these models cannot compensate for deviations from this symmetry.

$$x_{real} = \frac{x}{1 + \lambda_1 r^2 + \lambda_2 r^4 + \lambda_3 r^6 + ...} \tag{2.3}$$

$$y_{real} = \frac{y}{1 + \lambda_1 r^2 + \lambda_2 r^4 + \lambda_3 r^6 + ...} \tag{2.4}$$

2.2.2 Distortion Measurements

To approach this problem, an empirical approach is implemented. Rather than attempting to find the coefficients to fit a predetermined model to the data, the model is built by measuring distortion across the frame relative to fiducial features and interpolating between the measurement locations. This yields a mapping between the distorted and calibrated frame that is free of any symmetry assumptions. What does remain is an assumption of smoothness and continuity and that linear interpolation is sufficiently accurate over short distances in the frame.

The first step in empirically calibrating an image is the collection of distortion estimates across the image frame. There are several ways this can be performed. A popular method for established formula-based methods involves taking images of a checkerboard pattern positioned at a variety of distances and orientations with respect to the camera sensor. The vertices of the checkerboard pattern can be automatically detected in the image frame and then used to estimate the formula parameters [52]. The core requirement for any applied method is that imaged features can be matched with expected positions so that the distortion from truth can be estimated [53].

The method applied here leverages the data being collected during night sky observations to avoid requiring any sort of manual calibration. Rather, the stars in a given image act as the

sampled points. Existing star catalogs such as the Hipparcos star catalog have the angular positions of many of the brightest stars in the night sky to sub-arcsecond accuracy [54]. For an all-sky sensor such as the one used to verify this work, this is sufficient to determine the distortion from truth of an observed star to well below pixel resolution. This method for collecting calibration samples also allows the camera to be focused at infinity. Adjusting the focus of an optical system can alter the observed distortion, so by calibrating at the same setting used to collect data, the distortion map can accurately reflect the observed distortion during practical use.

The process for collecting the distortion samples starts by taking a series of images of the clear night sky. Any visible objects other than stars need to be masked out in order to avoid false detections that can severely impact the accuracy of the final calibration map. Next, the location of the stars within the frame are extracted. This can be done by taking a weighted centroid of every cluster of pixels above some threshold defined by the sensor's gain and exposure time. The goal in selecting this threshold to prevent noise from resulting in false positive star detections while keeping as many visible stars as possible.

The central region of an imaged frame tends to be the least distorted in the tested lenses. By cropping to keep only this central region, the approximate inertial coordinates of the center of the frame in right ascension and declination can be determined using a program such as Astrometry.net [55]. Astrometry.net also provides information about the angular resolution of the image which can be applied to build an initial guess for the inertial coordinates of every point in the frame. This initial guess for each observed star can be compared with the catalog stars true coordinates.

If the observed and catalog stars are sufficiently close to one another, an association algorithm such as a Hungarian algorithm can be applied to associate observed stars with their closest catalog analogues [56]. Should the distortion be too severe to correctly associate the two sets of stars, the image can be run through a functional calibration algorithm such as the Brown model to bring the observed star coordinates closer to the truth defined by the catalog.

Once the observed stars are correctly associated to the catalog stars, the distortion in both the horizontal and vertical image frame directions can be directly determined by the vector displacement

between the observed and catalog coordinates.

The basic method for associating applies the Hungarian algorithm to associate catalog and observed stars. First the images are pre-processed to subtract out the background and hot pixels. If the distortion in the optical system is severe, the Brown-Conrady method can be applied to remove much of the distortion. The images can then be processed though Astrometry.net to find the boresight inertial coordinates. With an inertial reference point, the visible region of the catalog can be projected onto the image. Next a cost function is built using the Euclidean distances in the image frame between each observed and catalog star. The cost matrix is fed into the Hungarian algorithm that then finds the optimal association between observed and catalog stars to minimize the total cost of the associations [56].

Algorithm 1 Basic association between observed and known stars.

$obsstars \leftarrow$ Weighted centroid coordinates of observed stars in the image frame.
$catstars \leftarrow$ Coordinates of catalog stars in the image frame.
for $i = 0, \quad i <$ length$(obsstars)$ **do**
    for $j = 0, \quad i <$ length$(catstars)$ **do**
        $costmatrix[i, j] \leftarrow$ norm$(obsstars[i] - catstars[j])$
$[obsorder, catorder] \leftarrow$ Hungarian$(costmatrix)$
$orderedobs \leftarrow obsstars[obsorder]$
$orderedcat \leftarrow catstars[catorder]$

2.2.3 Delaunay Based Calibration

While a series of images of the night sky can place distortion samples across the frame, the goal is to take an arbitrary point in the frame and determine the distortion from truth. To do this, the distortion in regions between observed stars needs to be interpolated based on its proximity to sampled points. This is done through the use of a Delaunay triangulation [50]. A Delaunay triangulation divides a two dimensional region into an irregular triangular tessellation with vertices defined by an arbitrary distribution of points.

There are many ways to divide a region into triangles. The Delaunay triangulation maximizes the minimum internal angle of all of the triangles [50]. This is done by ensuring that the

circumscribing circle for each triangle does not have any sampled points in its interior. The effect of this is to reduce the number of large obtuse triangles and ensure that the regions each triangle encloses is not too close to points that do not define a vertex. The Delaunay Triangulation is the straight line dual graph of the Voronoi Diagram which maps points on a plane to their nearest vertex [57].

The classic problem for the Delaunay triangulation is the task of interpolating the terrain of a map based on a discrete set of altitude measurements at arbitrary coordinates. The analogous nature of building a distortion map based on a distribution of measured points is why the Delaunay triangulation was chosen for this case. In terms of implementation, popular programming environment such as MATLAB and Python with SciPy have efficient native implementations of the Delaunay triangulation algorithm.

The Delaunay triangulation can then be used to build a pair of piece-wise linear surfaces where the x and y coordinates are defined by the locations of the sampled points in the image frame and the z coordinate is a linear interpolation of the distortion of the vertices. In this way, each triangle is effectively a plane segment describing the distortion in the region that it spans. Figures 2.1 and 2.2 show a uniformly selected random distribution of one thousand points across a frame and the corresponding Delaunay triangulation. To ensure that the full fame is covered by the triangulation, each corner is added as a sampled point.

In practice, after collecting matched sets of observed and catalog stars a point is added to each corner. The distortion associated with that point is set by the next closest point to the corresponding corner. The Delaunay triangulation creates a convex hull for the point cloud, so placing these four corner points ensures that the whole frame is tessellated. This tessellation and corresponding vertex distortion values is stored for every image corresponding to the same optical system. When a point is to be calibrated, the triangle that the point falls inside of is found for each tessellation. The distortion in the x direction is treated as the third coordinate of a three-dimensional space and the point is projected onto the plane defined by the triangle that contains it. Mathematically, the x direction distortion is given by equation (2.5) where $\vec{v}_{i,x}$ are the vertices

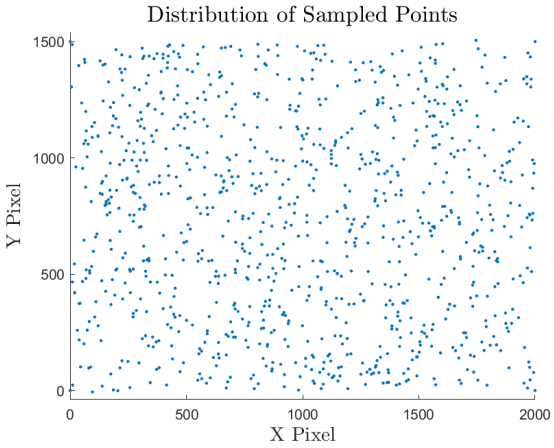

Figure 2.1: One thousand randomly selected points across an image frame.

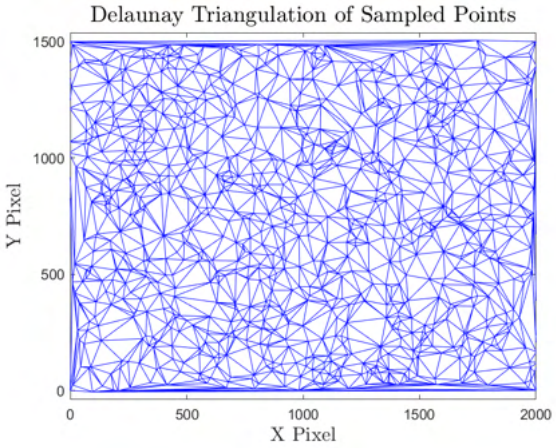

Figure 2.2: The Delaunay triangulation of the points in figure 2.1.

of the triangle, $\vec{p}_0$ is the location of the point of interest set on the $x-y$ plane, and $\hat{i}_3$ is the z direction unit vector. The y direction distortion is calculated using a similar equation shown in equation (2.6).

$$\Delta x = \frac{(\vec{v}_{1,x} - \vec{p}_0) \cdot [(\vec{v}_{2,x} - \vec{v}_{1,x}) \times (\vec{v}_{3,x} - \vec{v}_{1,x})]}{[(\vec{v}_{2,x} - \vec{v}_{1,x}) \times (\vec{v}_{3,x} - \vec{v}_{1,x})] \cdot \hat{i}_3} \tag{2.5}$$

$$\Delta y = \frac{(\vec{v}_{1,y} - \vec{p}_0) \cdot [(\vec{v}_{2,y} - \vec{v}_{1,y}) \times (\vec{v}_{3,y} - \vec{v}_{1,y})]}{[(\vec{v}_{2,y} - \vec{v}_{1,y}) \times (\vec{v}_{3,y} - \vec{v}_{1,y})] \cdot \hat{i}_3} \tag{2.6}$$

The results of these individual distortion estimates across all the tessellations can then be averaged together to provide a single estimate of the point's distortion. The effect of this averaging process is to smooth out the sharp edges at the borders of the triangles and dampen the impact of inaccurate vertex distortion measurements.

The Delaunay triangulation performs well at matching regions to their closest corresponding triplet of measured points. This approach does have its weaknesses. As shown in figure 2.2, the vast majority of the region's interior avoids long and narrow triangles. However, the region around the frame's edges are forced to consist of long and narrow triangles in order to fully cover the frame. As a result, points toward the center of these long and narrow triangles will be distorted by measurements relatively distant from them. Accordingly, distortion estimates from these regions will be unreliable. The other major weakness of this approach is a relatively low tolerance for errors. Incorrect associations between catalog and observed stars can have a large impact on the accuracy of nearby distortion estimates. Outlier detection and limiting the maximum association cost in the Hungarian algorithm can limit, but not completely negate the risk of accepting a false association.

2.2.4 Savitzky-Golay Calibration

An alternative to the Delaunay triangulation approach to turning measurements into a calibration map is the use of a two-dimensional Savitzky-Golay filter (SGF). The Savitzky-Golay filter is a local least-squares approximation method that, for a function of two variables, fits a series of

polynomials to local subsets of the data. Closely related to the Savitzky–Golay filter are variations of the locally estimated scatterplot smoothing (LOESS) method that differ from one another in the details of their implementation, but tend to produce similar results in practice [58].

So long as the underlying data is locally well-approximated by a low degree polynomial function, a SGF or LOESS method can be applied to find an piecewise smooth and continuous analytical function that estimates the observed data [58]. This resulting function can then be used as the calibration map by passing coordinates to be calibrated through the function.

Similar to the Delaunay triangulation, advanced numerical tools such as MATLAB and SciPy have Savitzky-Golay and LOESS tools to efficiently implement these functions into calibration algorithms. Unlike the Delaunay triangulation method where each image was individually triangulated and the piecewise maps were averaged together, it is more feasible to combine all of the calibration points into a single dataset and process the whole set with the SGF. This is because the filter handles the outlier discrimination itself, so it does not require an intermediary step to find and remove outliers as the Delaunay triangulation based method does.

The main benefit for switching to the Savitzky-Golay filter is that the algorithm is better able to detect and ignore outliers in the dataset and the endpoint behavior is less erratic. These are the major downsides to the Delaunay triangulation method, so the Savitzky-Golay filter can help build a better calibration map when there are a significant number of outliers in the data. The trade-off when applying a Savitzky-Golay filter, is that the method is less reactive to real deviations in the data, and the resulting map may smooth away small, but real, features that better describe the underlying behavior.

2.3 Results

2.3.1 Optical Observations with OmniSSA

The Raytheon Intelligence & Space VADeR Lab at the University of Colorado at Boulder has constructed a multi-sensor ultra-wide field-of-view optical system called the Omnidirectional Space

Situational Awareness (OmniSSA) system. The goal of this hardware project is to approach the problem of uninformed object detection and state estimation using four ultra-wide angle sensors. These sensors are composed of ZWO ASI1600MM cameras with Rokinon FE14M-E lenses in order to achieve 115 degree field-of-view images of the night sky. The fields of view of the four sensor overlap entirely. The output images can be used together in order to improve detection signal-to-noise ratio limits [23]. The system is designed for long term autonomous operation and remote access in order to passively observe the night sky and perform state estimation on a large set of observable space objects.

OmniSSA has been shown to be able to detect objects as dim as geostationary and geosynchronous space objects. The system is zenith pointing, so fast moving space objects such as those in low Earth orbit appear as streaks in the sensor frame for longer integration times.

OmniSSA is used to provide real data to observationally verify the calibration algorithm. The optical lenses used in this system induce distortion in the image relative to the expected projection of the sky. This provides a distorted target to apply the ultra-wide angle lens calibration. The large datasets from leaving OmniSSA to run for several hours while observing a significant portion of the sky provide ample targets for building the calibration maps using a well-distributed set of sample points.

Two approaches were taken to investigate the performance of these algorithms. The imaging system on OmniSSA has severe limb distortion and processing the observed data with Astrometry.net resulted in a very large number of false associations between observed and catalog stars. Accordingly, the Savitzky-Golay filter approach would be expected to perform better. Both methods were analyzed using data collected from OmniSSA's ultra-wide field-of-view imaging system. To better capture a case where the Delaunay triangulation method performed well, a simulated case that plays more to the Delaunay method's strengths is demonstrated as well.

2.3.2 Simulated Case

The Delaunay triangulation based method for calibrating optical systems is first tested using an artificially generated set of distortion samples. The purpose of this test is to measure the method's theoretical capabilities in the case of perfect assignment between observed and known coordinates. These tests look at the capability of varying numbers of image frames to accurately remove distortion. The tested images are set to have 2000 by 1500 pixels.

The tests are conducted by first choosing a distorting function to apply. The chosen function is a sine function applied in the radial direction with a magnitude of 15 pixels. The function is shown in equation (2.7) where δr is the change in radius with respect to the image center and r is the original point's radius from the image center. The true magnitude of the distortion in this mapping is shown in figure 2.3. Different numbers of simulated images are combined to produce estimated calibration maps each with 1000 virtual stars each. The resulting estimated distortion maps for one, five and twenty simulated images are shown in figures 2.4 through 2.6. The vector magnitude of the distortion is represented with the color of each point. To demonstrate how this breaks down into x-axis and y-axis distortion, figures 2.7 and 2.8 show the component distortions for the x-axis and y-axis respectively.

The single image calibration maps clearly show the underlying triangles from the Delaunay triangulation. As more images are averaged together, the pattern disappears. The final remaining difference between the twenty image result and the underlying truth is the edge distortion as discussed in the theory section. The radial sinusoidal curve was chosen because it is a case that would be a poor fit for the standard Brown and division models and it shows fairly dramatic shifts between high and low distortion areas.

$$\delta r = 15 \sin\left(\frac{r}{160}\right) \tag{2.7}$$

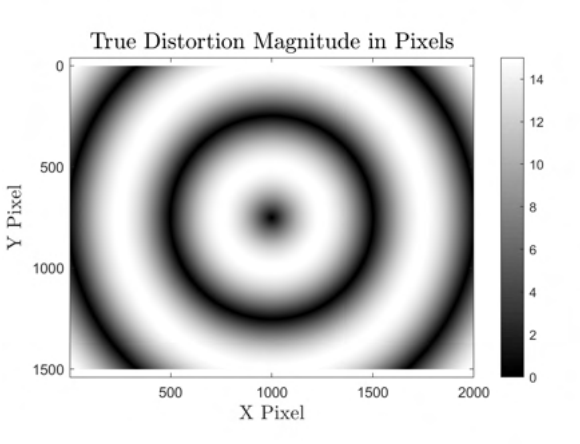

Figure 2.3: True distortion magnitude in pixels for the simulated case.

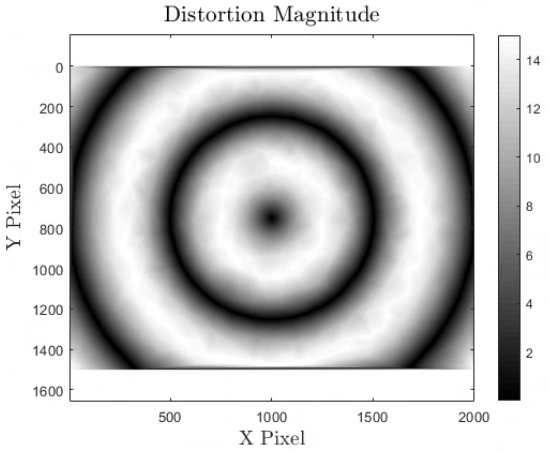

Figure 2.4: Estimated distortion for a map made from a single simulated image.

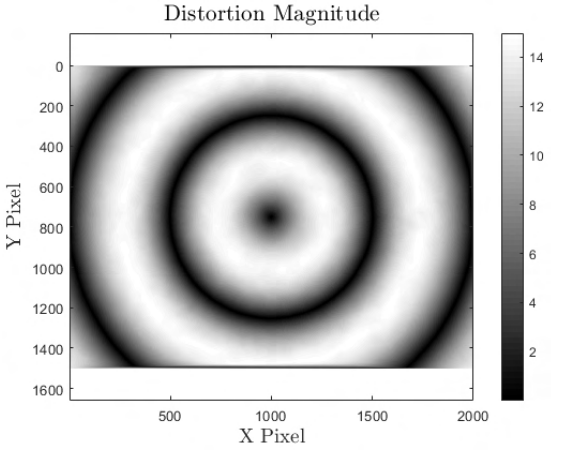

Figure 2.5: Estimated distortion for a map made from 5 simulated images.

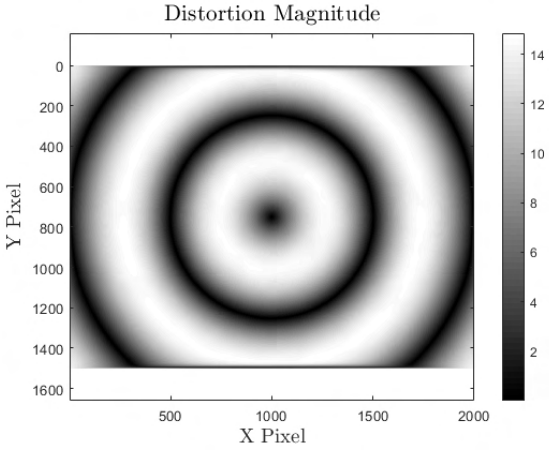

Figure 2.6: Estimated distortion for a map made from 20 simulated images.

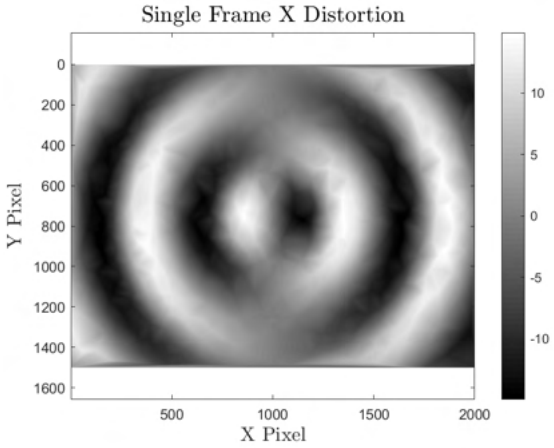

Figure 2.7: Estimated x-axis distortion for a map made from a single simulated image.

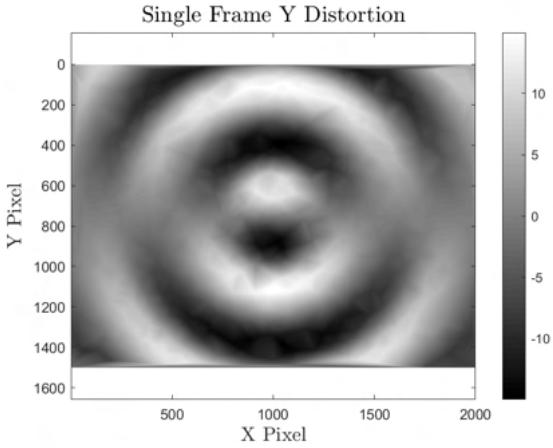

Figure 2.8: Estimated y-axis distortion for a map made from a single simulated image.

2.3.3 Empirical Verification

To empirically verify the methods, calibration maps were generated with sets of images and another set of images taken with the same camera was used to extract the error residuals. Associations were made using the star identification feature of Astrometry.net. The larger set of several thousand associations that went in to building the map were fed directly from the results of Astrometry.net. For both cases, the Brown model calibration used by Astrometry.net was applied first to pre-calibrate the image and help reduce the initial residuals before applying the empirical calibration. The data used to measure the final residuals was checked to confirm that the associations were correct so that the results could be appropriately compared to truth.

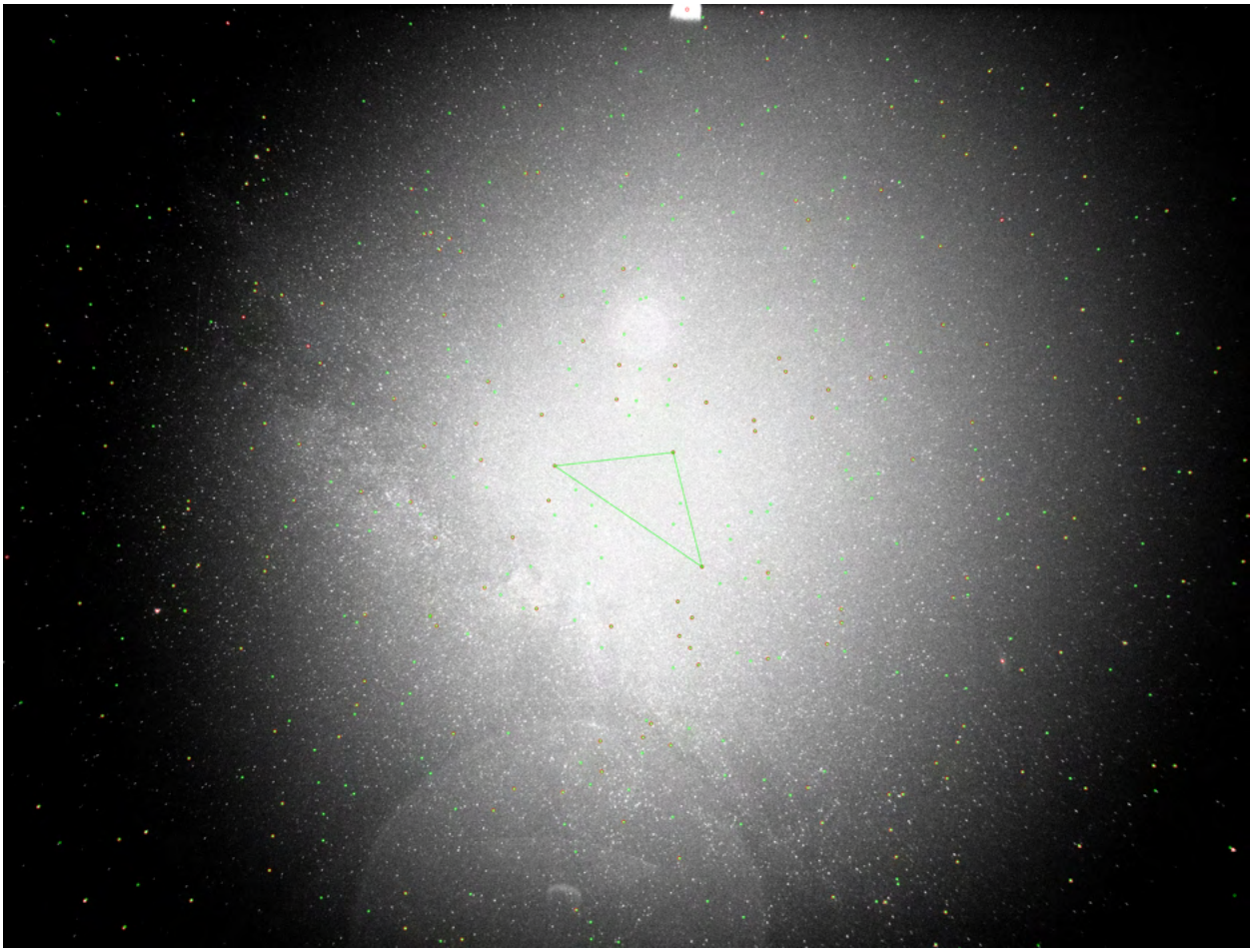

Figure 2.9: Stars detected (red) and expected (green) by the Astrometry.net software.

Both calibration methods take an input of a series of images taken by OmniSSA. The two cases are tested on distinct yet similar datasets using two of the cameras on OmniSSA. For the Delaunay triangulation case, the distortion magnitude across the image frame for the average over fifty triangulations is shown in figure 2.10. There were a large number of false associations in the dataset. They were concentrated further away from the central region, and that is where the most significant non-linearities appear to lie in the resulting calibration map. A more numerical analysis of the method's accuracy is found by finding the residuals between known pairs of catalog and observed stars in images taken by this camera.

The residual distance between the expected and observed locations can be found for each case both with and without the correction provided by the calibration map. The cumulative distribution function of the residual between expected and observed star location before and after calibration is shown in figure 2.11. Each pixel is 89 arcseconds across as designated by the tic marks along the x-axis. To reduce the number of outliers in the data, associations that cost over 0.09 degrees were removed as likely false associations. This represents a slight improvement in the residual lengths for known star associations. The portion of stars within 1 pixel of their expected location rose from 27% to 51% while the portion within 2 pixels of their expected location rose from 87% to 96%.

A similar dataset was then processed using the Savitzky-Golay filter approach. As described in the theory section, this used a combined dataset from all of the data images to build the resulting map. This method does more to smooth the resulting data, and the resulting calibration map shows this. Figure 2.12 shows the magnitude of distortions in the distortion map generated with the Savitzky-Golay filter. Repeating the same technique as the Delaunay triangulation method to find the residuals, the semi-log plot for the cumulative distribution function is shown in figure 2.13. In this case, 81% of the calibrated residuals fell within one pixel of the expected value, and 93% fell within 2 pixels of their expected value.

The Savitzky-Golay filter can be adjusted in terms of how large of an area the filter includes in each local component. With the implemented algorithm, this is expressed in terms of the fraction of the overall dataset that is included for each local fit. Several cases are run for different values.

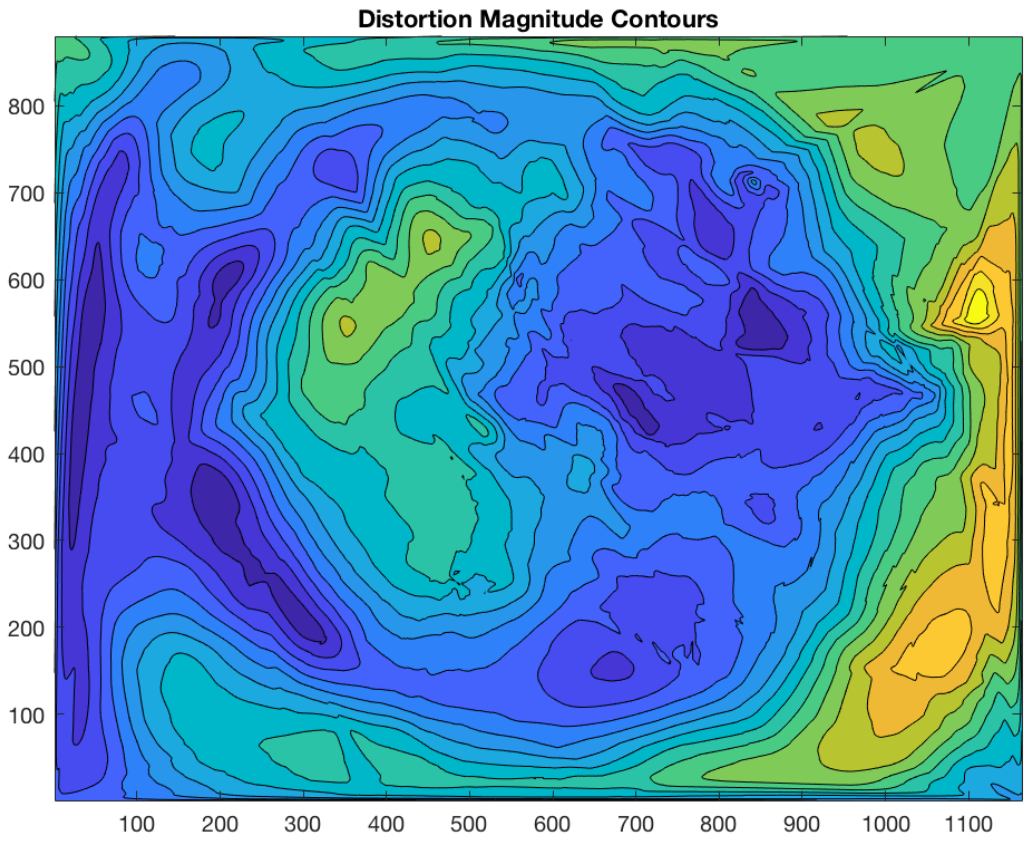

Figure 2.10: Estimated distortion on a fifty image OmniSSA dataset processed by the Delaunay triangulation.

The results are provided in figures 2.14 and 2.15. As the datasets were very large with data from 105 images, smaller spans provided slightly better results. The best results were seen with one percent of the data in each local fit with the accuracy of the results falling off as the size of the span increased. For all cases between sixty and seventy percent of test points had residuals less than one pixel. The one percent span performed the best with sixty nine percent of residuals less than one pixel. The best performing span changed around those with residual of two pixels. Span of twenty percent performed best at that point with eighty seven percent of residuals less than two pixels. Smaller spans follow the underlying data more closely but are more sensitive to the influence of outliers. The change in best performing spans illustrates this behavior.

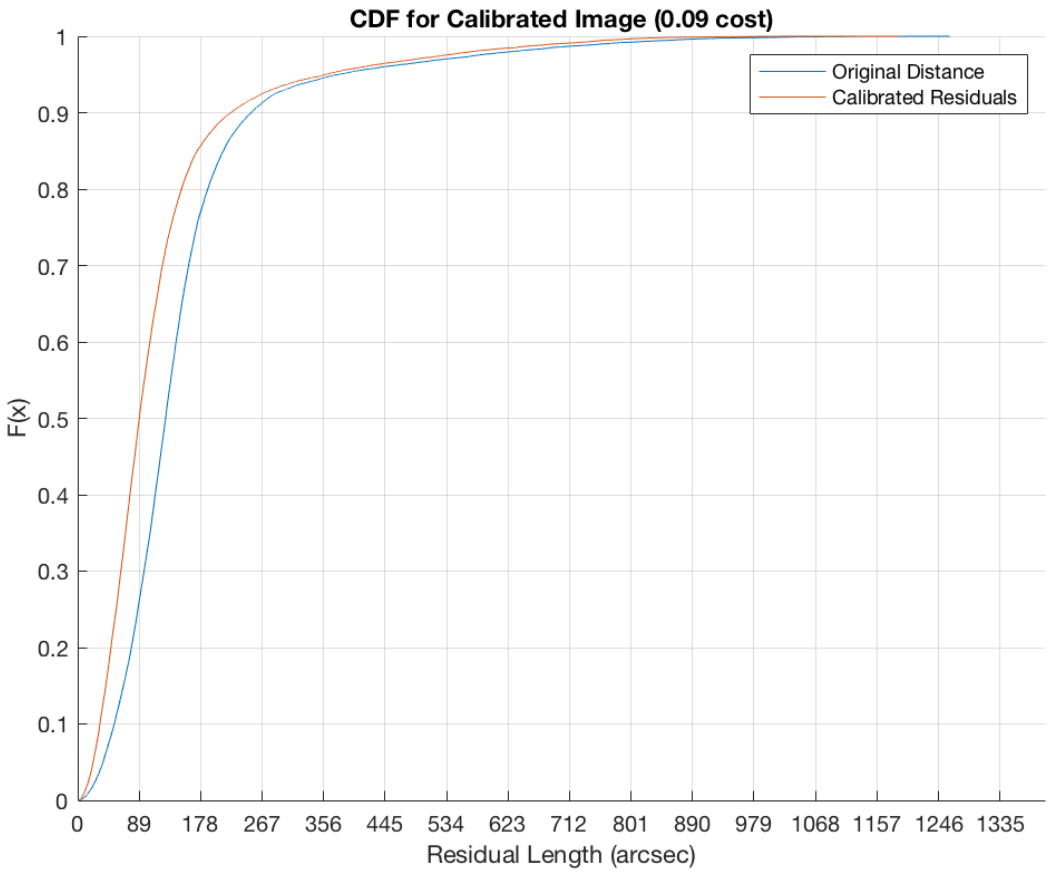

Figure 2.11: CDF for Delaunay triangulation calibration of the first OmniSSA sensor.

The one percent span performed much better than any other case for the period that it was the best performing method. When it lagged behind, it did not lag behind by much. It swapped back to best performing after a very short time. So for this dataset, the one percent span was the appropriate setting for minimizing residuals.

2.4 Chapter Summary

An empirical approach to ultra-wide field-of-view optical system calibration was developed and discussed. This method makes no assumptions regarding the underlying symmetry of the true distortion. Rather, the only core assumptions is that the error at a point is consistent from image to image and the error itself is smooth and continuous across the image area. With these assumptions,

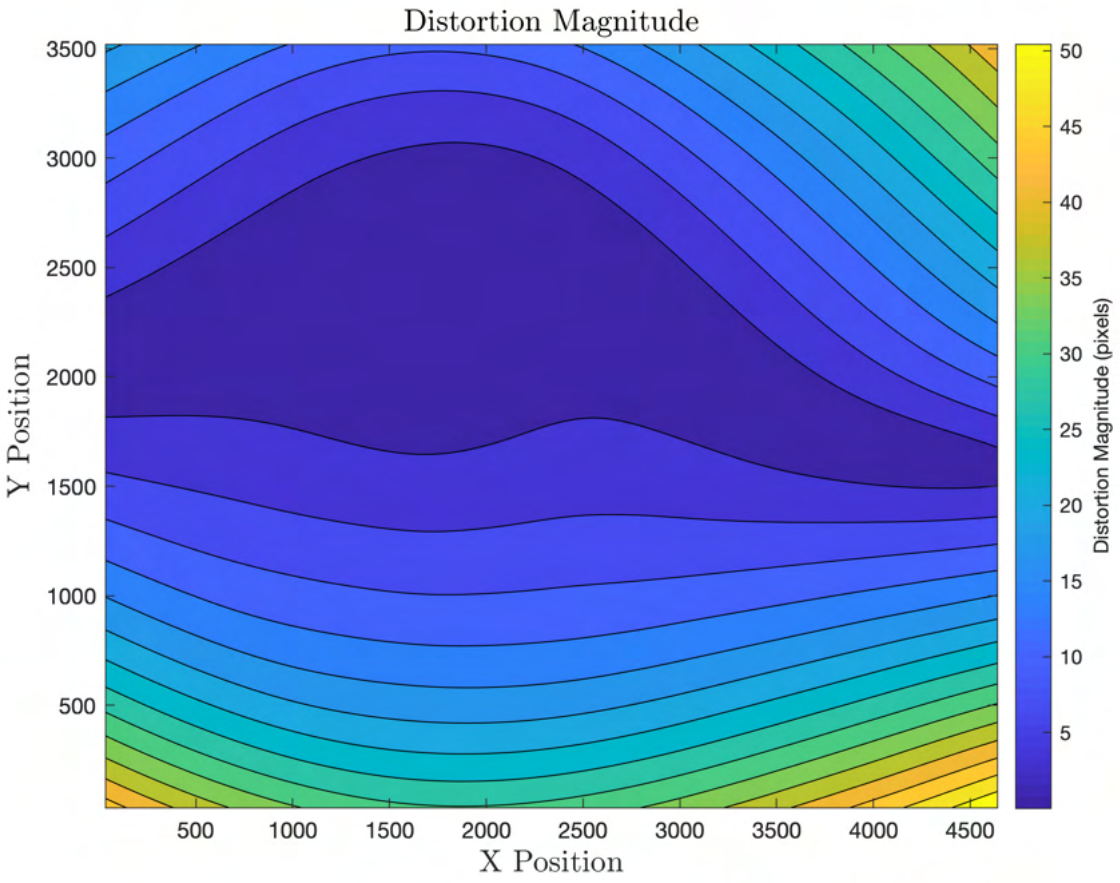

Figure 2.12: Estimated distortion on a fifty image OmniSSA dataset processed by the SGF.

the empirical approach was developed. The process relies on extracting calibration points based on the expected and observed locations of stars in night-sky images taken by the optical system. These are then used as reference points for building a continuous surface. Two different methods for constructing this surface were discussed. The first involved building a piecewise continuous set of triangular plane segments that spanned the image frame. The Delaunay triangulation was applied to optimally assign regions to closest neighbor sample points. In order to overcome shortcomings in some cases where there are high numbers of false associations, a second method was proposed using a Savitzky-Golay filter to smooth the resulting surface best estimated from the calibration points.

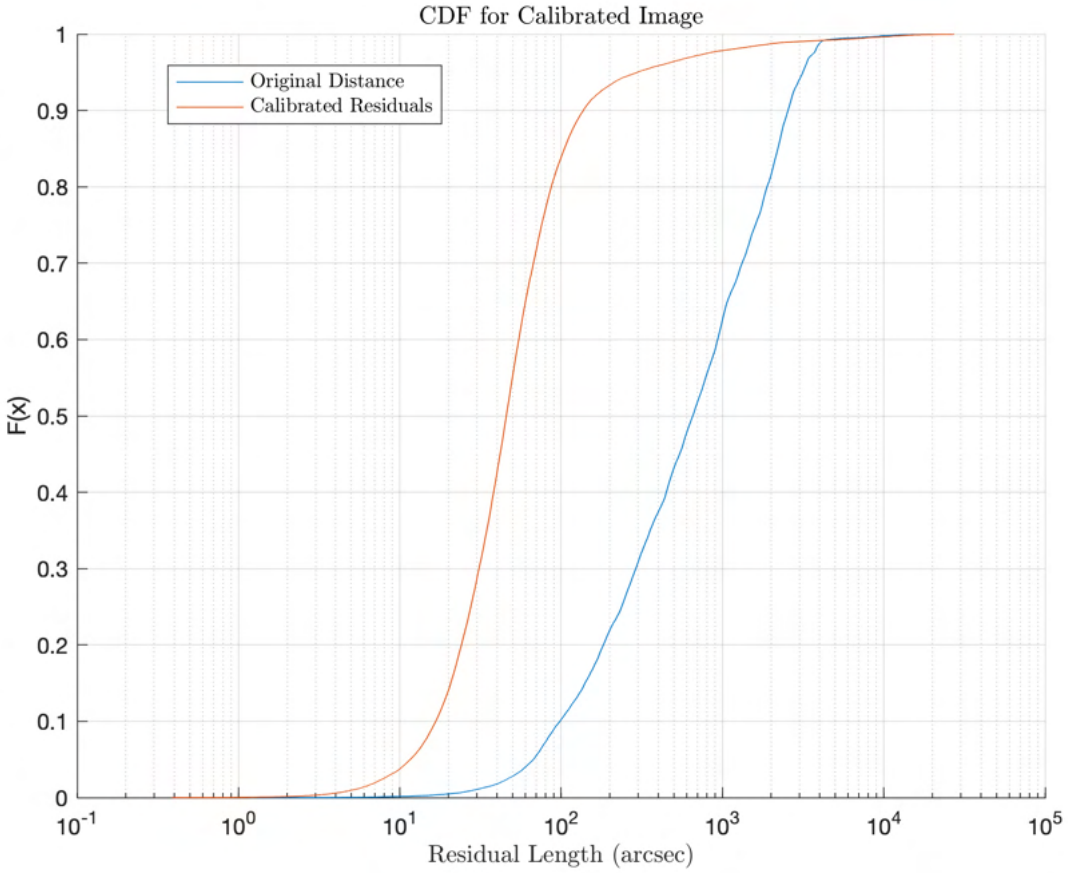

Figure 2.13: Semi-log plot of the pre and post SGF calibration error residuals between estimated and catalog star locations.

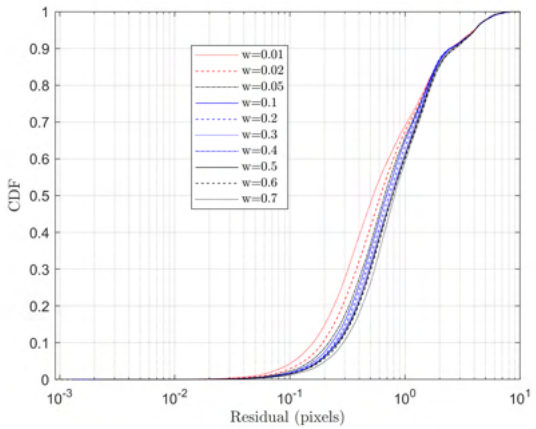

Figure 2.14: CDF of residuals for Savitzky-Golay fits with spans from 0.01 to 0.7.

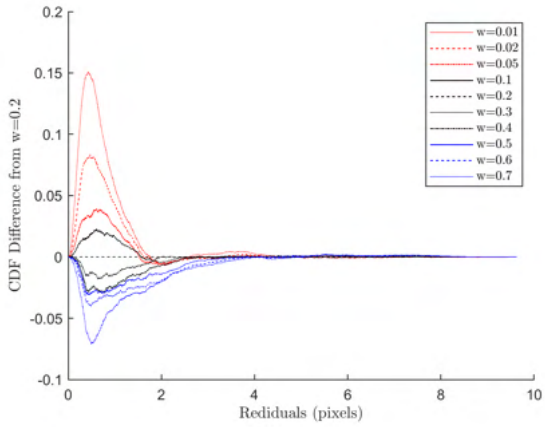

Figure 2.15: Difference in the CDF plots for each case with the $w = 0.2$ case. Larger values indicate more of the sample points with lower residuals.

Chapter 3

Optimal Extraction of Photometric Light Curves

3.1 Streaking Objects and Light Curves

A moving space object can appear as a streak in an image from a non-tracking optical system. Compared to a tracking optical system, the photometric signal-to-noise ratio (SNR) for a streak will be lower as the observed signal is spread out over a larger region of the imaging frame and each individual pixel will be closer to the background noise floor of the image [59]. While this loss of SNR is undesirable in many applications, in others this may not pose a problem or not be possible to counteract. In the first case, the signal of the resulting streak may already be sufficiently high above the noise floor to remain clearly observable for the intended application [28]. In the second case, the trajectory of the target object may not be known in advance, such as the case with target acquisition applications and unexpected maneuvers [60]. If the target object remains observable as a streak, there is additional information that an optical system can extract from relating to the object's photometry that are not observable for a perfectly tracked object.

As an object streaks across a frame while the sensor integrates the incident light, the intensity of the light received by the sensor may change over the integration period [61]. This can reflect changes in the albedo of the object as it rotates relative to the original light source and the observer, or it could reflect the translational motion of the object as it passes through the Earth's penumbra or umbra [62]. In either case, the photometric information contained in the streak for objects moving with respect to the imaging frame can provide additional insight into the transnational or attitude state of the object [15]. Accordingly, extracting this information and determining the certainty

regarding the extracted values are valuable tools for better understanding the target object and its state.

Having identified the additional information available contained within the streak's photometry, there remains the question of how to go about extracting the information and determining the correct amount of confidence to have in the corresponding estimate. Since the intensity of the incident light can vary continuously, the true photometry is some continuous function of time [63]. This light from the target object passes through the physical optical system of the observing instrument as well as any atmosphere between the observer and target. Light from a point source target will be distorted to a point spread function as it passes through this system until it reaches the sensor. The image captured by the sensor will then reflect the integrated motion of this point spread function as it is smeared across the sensor [64].

While estimating the continuous function that defines the observed object's photometry, there is a limit to the amount of meaningful information that can be extracted from a given streak due to the combined impact of this point spread function with the inherent measurement uncertainty for the value reported by each pixel in the sensor [65]. This measurement uncertainty is the result of the combined impact of both the observed background light as well as the physical limitations of the real sensor. Extracting the photometry of a streaking object requires first finding a model to represent the underlying time-dependent photometry function and then finding an estimated function based on the model that appropriately represents the available and observable information contained in the imaged streak. Additionally, a complete estimate should be paired with an estimate of the uncertainty in the reported values. Both the estimated model as well as the uncertainty are necessary to fully report the results.

Current methods for extracting the photometry of streaking object largely ignore the problem of finding the uncertainty in the resulting estimates [16][17]. Some approach the problem by manually cutting the the image integration time so as to break what would have been a longer streak into a series of separate images and a single estimate of the signal intensity is found for each image in order to find the change with time [16]. Another existing approach is to partition the observed

streak into a series of boxes and find the corresponding photometry of the streak within each of these boxes. However, this approach does not offer any method for finding an appropriate number of boxes. It also does nothing to account for the point spread function bleeding into multiple of these boxes [17]. The current state of the art takes an ad hoc approach to choosing how to break down the streak into parts and does not contend with the issue of finding the uncertainty of the resulting light curve so as to provide some measure of the confidence in the provided estimates.

A novel method for extracting the photometry of straight-line space object streaks is investigated to determine an optimal approach and provide a method for estimating uncertainty in the photometry. The proposed method models the photometry function as a stairstep function where the streak is broken down into a series of streak segments that are each constant and evenly spaced across the streak's length. The photometric intensity and corresponding uncertainty are found with a non-linear least squares regression. The core innovation offered with this method is an approach for finding the optimal number of streak segments to break the full streak down into in order to extract all observable information without attempting to estimate any unobservable states by breaking the streak into too many segments.

Two methods for determining the appropriate number of streak segments are proposed. The first is an analytically derived equation for measuring how information optimal a trial value for the number of streak segments. This analytical formula is particularly computationally efficient to implement, however, it is based on several core assumptions regarding the geometry of the streak within the image plane. This means that the method is not universally applicable, as when these key assumptions regarding the streak break down, the value that the formula produces may not reflect the true optimal value for the number of streak segments. To overcome this, the second method is a computational approach that is less efficient in terms of the computational resources it requires, but it does not rely on any assumptions about the streak's geometry and is therefore universally applicable for any given straight-line streak. This includes streaks generated by slitless spectroscopy where the photometric information encodes the frequency distribution of the received light. An optimal breakdown of spectra from slitless spectroscopy will allow the maximum resolution

spectrum to be extracted for an object.

The proposed methods are validated against both simulated and observed data. These cases demonstrate how all of the available information is extracted by an appropriate number of streak segments. Simulated cases are also used to verify that the different proposed methods converge to the same results when the core assumptions for the methods are met. This demonstrates how under appropriate conditions, an efficient method can be used to determine the optimal method for breaking down a streak and estimating the photometry along its length.

3.2 Theory and Models

3.2.1 Constructing the Problem

An observed stationary object will appear in an image as a bivariate Lorentzian distribution [66]. This can be locally approximated as a bivariate Gaussian distribution in order to take advantage of the Gaussian function's superposition properties [67]. If the camera system has been calibrated, the bivariate Gaussian should by radially symmetric and have a covariance matrix that can be represented as equation (3.1) where σ is the standard deviation based on the light wavelength and optical system parameters, and $I_{2\times2}$ is the two by two identity matrix.

$$C = \sigma^2 I_{2\times2} \tag{3.1}$$

First, assume that the photometric intensity can be modeled as some continuous function of time, $B(t)$, over the duration of the image's integration time. From the perspective of a single pixel, a Gaussian curve slides over it, and the total intensity is integrated over the motion. For a rectangular pixel with horizontal and vertical limits defined by x_1, x_2, y_1, and y_2, the total intensity observed by the pixel is given by equation (3.2). In this equation, the μ terms are the time dependent coordinates of the point spread function (PSF) mean and the image integration time goes from t_0 to t_1.

$$I = \int_{t_0}^{t_1} \int_{x_1}^{x_2} \int_{y_1}^{y_2} \frac{B(t)}{2\pi\sigma^2} \exp\left(\frac{-1}{2} \begin{bmatrix} x - \mu_x(t) \\ y - \mu_y(t) \end{bmatrix}^T \sigma^{-2} I_{2\times 2} \begin{bmatrix} x - \mu_x(t) \\ y - \mu_y(t) \end{bmatrix} \right) dydxdt \quad (3.2)$$

To visually demonstrate how the streak is broken down, figure 3.1a shows contours for the first three standard deviations for a symmetric bivariate Gaussian superimposed on a grid to represent the pixels of a sensor. This point spread function slides across the image frame as the source moves through the field of view. The result is that the integrated light shows up as the point spread function smeared across the plane. This is represented in figure 3.1b where the point spread function from figure 3.1a was used to define the distribution of randomly generated virtual photons as the point spread function slid across the frame. The image shows where each of these photons struck the sensor. Figure 3.1c demonstrates what the camera would see as a result. Each pixel takes a value corresponding to the number of photons to land within its borders. The sensor is effectively a two-dimensional histogram of the incident light with bins defined by the pixel locations.

The only information about time available for this image is the shutter open and shutter close time and the total integration time given by their difference. The intensity can be rewritten independently of time by replacing the point spread function mean as a function of time with a point spread function mean as a function of position along the streak, r. The intensity term, $B(t)$ can be exchanged for one dependent on the position along the streak $B(r)$, and the time integral that represents the time dependent motion of the point spread function can be replaced with an integral that directly tracks the translational position of the mean. Rewriting equation (3.2) in terms of position along the streak, r, yields equation (3.3). There the total length of the streak is given by l.

$$I = \int_{0}^{l} \int_{x_1}^{x_2} \int_{y_1}^{y_2} \frac{B(r)}{2\pi\sigma^2} \exp\left(\frac{-1}{2} \begin{bmatrix} x - \mu_x(r) \\ y - \mu_y(r) \end{bmatrix}^T \sigma^{-2} I_{2\times 2} \begin{bmatrix} x - \mu_x(r) \\ y - \mu_y(r) \end{bmatrix} \right) dydxdr \quad (3.3)$$

The $B(r)$ term only depends on r and can be pulled out of the position integrals yielding equation (3.4).

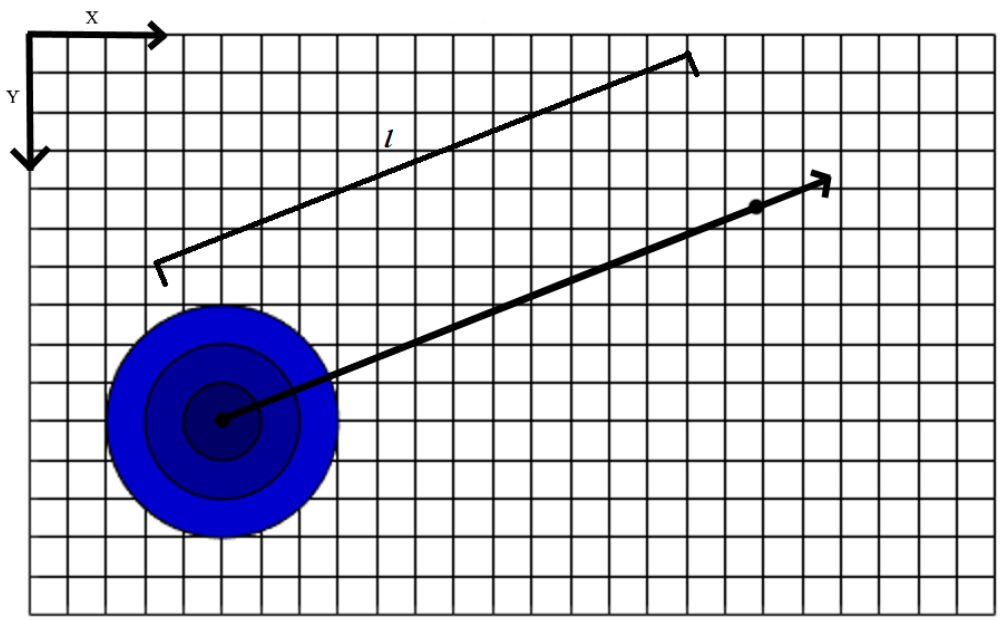

Figure 3.1a: One, two, and three sigma contours for a simulated bivariate Gaussian PSF.

$$I = \int_0^l B(r) \int_{x_1}^{x_2} \int_{y_1}^{y_2} \frac{1}{2\pi\sigma^2} \exp\left(\frac{-1}{2} \begin{bmatrix} x-\mu_x(r) \\ y-\mu_y(r) \end{bmatrix}^T \sigma^{-2} I_{2\times 2} \begin{bmatrix} x-\mu_x(r) \\ y-\mu_y(r) \end{bmatrix} \right) dydxdr \quad (3.4)$$

The $B(r)$ function is the photometric intensity function that is to be estimated from the the streak. The function can be approximated with a stair step function $B^*(r,n)$. This function is a piece-wise constant function with n steps. Each step is the mean value of $B(r)$ over the length of the step. The step length is simply the total streak length divided by the value of n. As n approaches infinity, $B^*(r,n)$ approaches $B(r)$. This is shown mathematically in equation (3.6). Due to the limits imposed by noise, there is some minimum value of n that allows $B^*(r,n)$ to model $B(r)$ well enough that the difference between a modeled streak based on $B^*(r,n)$ and the observed streak is

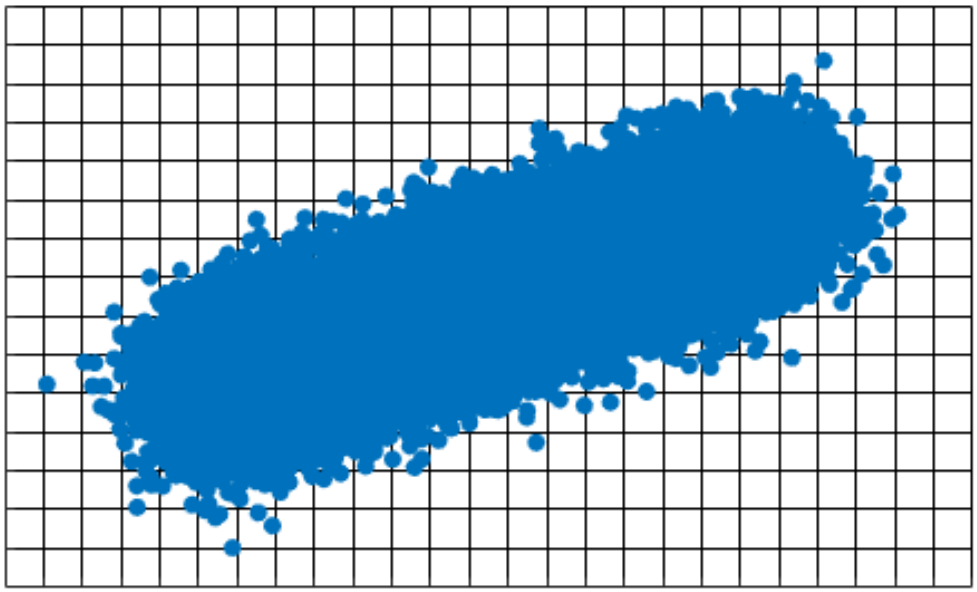

Figure 3.1b: Simulated distribution of photons on a virtual sensor represented as a grid.

additive white Gaussian noise. At that point, all observable information has been extracted from the image [68].

$$r_i < r < r_{i+1}$$

$$B^*(r,n) = \frac{1}{r_{i+1} - t_i} \int_{r_i}^{r_{i+1}} B(r)dr \tag{3.5}$$

$$B^*(r,\infty) = \frac{1}{\delta r} \int_{r-\delta r/2}^{r+\delta r/2} B(r)dr = B(r) \tag{3.6}$$

The input measurement for the streak photometry extraction scenario is a single picture taken by a camera with some resolution. The extracted streak from the frame will only occupy a fraction of the full frame. After identifying the streak, the subsection of the frame fully containing

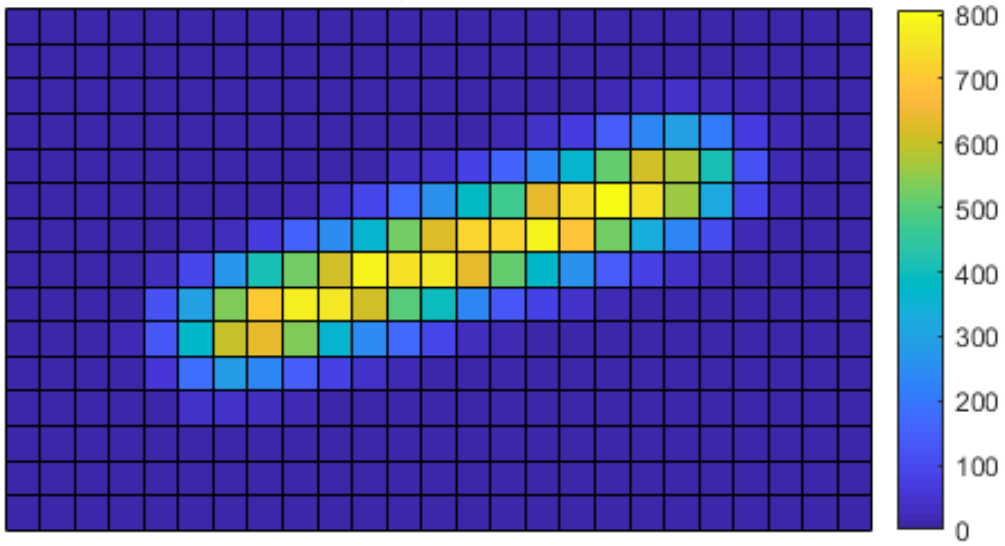

Figure 3.1c: Integrated result of the virtual photon distribution from figure 3.1b. Generated with PSF standard deviation of 1 pixel and length 15.8 pixels

the streak can be isolated and modeled as some $N \times M$ array of pixels. Working with this subframe alone reduces the amount of data that needs to be processed through the algorithm and reduces the computational load of estimating the streak's photometry. This frame reduction is important for building an efficient implementation of the algorithm as the calculation speed scales with the number of pixels in the frame.

The $N \times M$ subframe is built out of $N_P = N \times M$ pixels. Each pixel value is a measurement of the light incident on that portion of the sensor and is stored as an integer between zero and the bit depth of the sensor minus one inclusive. Each pixel measurement can then be treated as a single measurement in a set of N_P simultaneous measurements. By reshaping the $N \times M$ matrix of

pixel values into a single column of N_p values, the image data can be represented as a measurement vector, $\vec{y}$.

Equation (3.4) is adjusted to replace the true photometry, $B(r)$ with the estimated stair-step photometry $B^*(r,n)$ as shown in equation (3.7). The outermost integral can then be separated into the sum of n integrals with each one integrating over $\frac{1}{n}$ of the total imaged streak. By separating out the streak segments this way, the photometry $B^*(r,n)$ is constant within each integral. This constant can be relabeled as B_i^* for the i^th streak segment then pulled out of the integral entirely resulting in a total pixel intensity equal to a linear combination of integrals. This is expressed in equation (3.8). Each term in the summation corresponds to a segment of the total observed streak.

$$I = \int_0^l B^*(r,n) \int_{x_1}^{x_2} \int_{y_1}^{y_2} \frac{1}{2\pi\sigma^2} \exp\left(\frac{-1}{2} \begin{bmatrix} x-\mu_x(r) \\ y-\mu_y(r) \end{bmatrix}^T \sigma^{-2} I_{2\times 2} \begin{bmatrix} x-\mu_x(r) \\ y-\mu_y(r) \end{bmatrix} \right) dydxdr \quad (3.7)$$

$$I = \sum_{i=1}^{n} B_i^* \int_{r_i}^{r_{i+1}} \int_{x_1}^{x_2} \int_{y_1}^{y_2} \frac{1}{2\pi\sigma^2} \exp\left(\frac{-1}{2} \begin{bmatrix} x-\mu_x(r) \\ y-\mu_y(r) \end{bmatrix}^T \sigma^{-2} I_{2\times 2} \begin{bmatrix} x-\mu_x(r) \\ y-\mu_y(r) \end{bmatrix} \right) dydxdr \quad (3.8)$$

The values for the streak segment intensities making up B_i^* can be arranged into a $n \times 1$ column vector, $\vec{x}$. Now there is a measurement vector, $\vec{y}$, corresponding to every observed pixel intensity as well as a state vector, $\vec{x}$. Equation 3.8 relates these two quantities for a single pixel. As the values of B_i^* do not change for different pixels, this function relates every pixel value in the image to the state vector, $\vec{x}$. Since the measurements are linear combinations of the integral values and the state elements, the two can be separated to form a single linear equation shown in equation (3.9). H is the $N_P \times n$ measurement matrix composed of all of the integral terms for each streak segment's impact on each pixel.

$$\vec{y} = H\vec{x} \quad (3.9)$$

The terms in the H matrix only depend on the the geometry of the streak which can be determined as soon as the steak is identified and localized. The first of these terms is the standard

deviation of the point spread function, σ, which should depend purely on the optical system for observing point sources. Next are the point spread function means in each dimension, $\mu_x(r)$ and $\mu_y(r)$, which can be estimated based on the streak endpoints and applying some assumption regarding the point spread function motion. Equation (3.9) can then be solved as a linear least squares problem. The covariance matrix for the problem, P can be estimated based on the measurement covariance for each pixel, an $N_P \times N_P$ matrix, R. The equation for the covariance matrix is given in equation (3.10).

$$P = (H^T R^{-1} H)^{-1} \tag{3.10}$$

It is assumed that the point spread function moves at a constant rate over the duration of the image integration. However, if a more accurate model of the point spread function's motion is available, that can be applied after finding the photometry across the streak in order to find the time-dependent photometry. Under the constant velocity assumption, translating from translational to time-dependent photometry can be done by dividing the distance along the streak by the velocity of the point spread function.

The remaining problem is to find the appropriate value of n that represents the minimum number of states required to model all observable information in a given streak. To approach that problem, first a method for determining what defines an optimal value for n must be articulated. n sets the number of columns for the measurement matrix, H. Since this system uses only a single measurement vector, the observability gramian for the system is simply $H^T H$ which is a $n \times n$ matrix.

The determinant of this gramian quantifies the amount of information modeled by the matrix. This is sometimes referred to as d-optimality and when the matrix is an observability gramian, the d-optimality of the matrix is maximized when the observable information is maximized [69]. The determinant of the observability gramian will increase as the number of states increase. However, once the number of states is high enough such that the system is overdetermined, the H matrix

will become linearly dependent and, theoretically, the observability gramian will approach zero. In practice, the observability gramian does not perfectly approach zero due to the observed data not perfectly matching the model, however, after exceeding the optimal value for n, the determinant will decrease. The optimal value for n can then be found by finding the value that maximizes the determinant of the observability gramian.

Two solutions for this problem are proposed. The first is an analytical approach to the problem that applies a series of assumptions to manipulate the model until deriving an equation that quickly finds the expected observability gramian determinant for a given value of n and the length of the streak in pixels. Determining an optimal value for n is as simple as iterating up through the possible values until reaching the streak length and taking the value of n that returns the highest determinant. This method is very fast as it only requires the calculation of a single function for each trial value of n. The second is a computational algorithm that finds the optimal value of n for any given streak. Depending on the implementation, it either directly computes the integrals of the streak for each pixel or it uses one of several suggested approximation methods. This is effectively a brute-force approach to the problem and is far less computationally efficient. However, the method does not make any assumptions about the streak beyond the basic model and is applicable in many situations that the analytical approach fails such as short streaks or streaks where the pixels are not small relative to the streaks.

3.2.2 Analytical Approach

Starting with no assumptions simplifying the approach, the point spread function is taken as a bivariate Lorentzian distribution [66]. Replacing the simplified Gaussian assumption used in equation (3.2) with the bivariate Lorentzian distribution yields equation (3.11). There the scale parameter, γ, determines the shape and size of the distribution as the Lorentzian distribution does not have a defined mean or standard deviation [70]. A similar method as shown in the previous section can be applied to rewrite equation (3.11) in terms of a stairstep function and remove the time dependence. The result is expressed in equation (3.12).

$$I = \int_{t_0}^{t_1} \int_{x_1}^{x_2} \int_{y_1}^{y_2} \frac{B(t)\gamma}{2\pi((x-\mu_x(t))^2+(y-\mu_y(t))^2+\gamma^2)^{3/2}} dydxdt \quad (3.11)$$

$$I = \sum_{i=1}^{n} B_i^* \int_{r_i}^{r_{i+1}} \int_{x_1}^{x_2} \int_{y_1}^{y_2} \frac{\gamma}{2\pi((x-\mu_x(r))^2+(y-\mu_y(r))^2+\gamma^2)^{3/2}} dydxdr \quad (3.12)$$

Equation (3.12) provides the intensity observed by a single pixel written as a sum of the stairstep intensities. The elements of the summation can be broken out to form the rows of the measurement matrix, H. Each pixel is a row and the columns correspond to the streak segment contribution to the pixel's total observed intensity. The observability gramian is $H^T H$. The elements are the element-wise products of the columns. Accordingly, the diagonal is the product of a single streak segment's contribution to all pixels squared and then summed. The off-diagonal elements are the products of the contribution of different streak segments. An inefficient, but accurate method for finding the optimal number of states is to use trial values of n and build the observability gramian by calculating the integrals for each pixel. The determinant of the observability gramian can then be used to compare the gramians to one another to find the value of n that is d-optimal and therefore information optimal. Once again, this method is particularly computationally inefficient. By applying several reasonable assumptions regarding the shape of the streak, a much more fast and efficient approach is possible.

3.2.3 Simplified Analytical Approach

By applying several key assumptions, there is an analytical solution to the optimal number of states problem. The key assumptions that enable the analytical solution are as follows.

(1) Pixels are small relative to both the streak length and width. $x_2 - x_1 \ll l, x_2 - x_1 \ll 2\sigma, y_2 - y_1 \ll l, y_2 - y_1 \ll 2\sigma$

(2) The streak is much longer than its point spread function. $2\sigma \ll l$

(3) Any single pixel's incident intensity can be well approximated as the sum of the influence of at most two streak segments.

The measurement matrix, H, breaks the streak into n columns with each column representing the impact of a single streak segment on the full image. The streak segment is itself associated with a length of $\frac{l}{n}$ within the full streak length of l. The streak segment is generated only by the integrated intensity of the observed object from some r_i to $r_{i+1} = r_i + \frac{l}{n}$. Any region sufficiently far from the center-line of the streak segment during the associated duration will receive a very low exposure. A given pixel can receive a total intensity defined by one of three cases. Pixels far from the full streak will receive a very low intensity of light and report a low value. Pixels that lie towards the middle of streak segment will report a value dominated by the contribution of a single streak segment. Pixels towards the edge between two streak segments will report an intensity defined primarily by the contribution of two streak segments.

The elements of the observability gramian, $H^T H$, can be defined based on these pixel categories. The diagonal values are the sum of the square values of all pixels associated with a single streak segment. This can be found by generating a model streak segment, and adding up the square of each pixel. The off-diagonal elements are the sum of the products of the value of each pixel at two different times associated with different streak segments. By applying assumption three, it can be approximated that of the off-diagonal elements only the matrix elements representing adjacent streak segments are significantly above zero. All other elements correspond to the pixel-wise products of distant streak segments such that each result is the product of at least one very small value. The sum of all these products would then be very small itself. The result is that the observability gramian can be approximated as a tri-diagonal matrix.

As established in the previous section, the optimal value for n is the value that results in the maximum determinant for this matrix. For a tri-diagonal matrix, the determinant can be found as the result of the recursive function expressed in equation (3.13) [71]. There, the a terms are the values of the diagonal elements while the b terms are the values of the off-diagonal elements. The

subscript associated with each of these elements corresponds to the index along the diagonal. Since the observability matrix is the result of a matrix multiplied by its own transpose, the resulting matrix is symmetric.

$$f_n = a_n f_{n-1} - b_{n-1}^2 f_{n-2} \tag{3.13}$$

With the recursive form established, the values for the a and b terms are all that remain for finding an equation for the determinant for a given value of n. The diagonal elements of the matrix are the sum of the square pixel values. For small pixels as established in the first assumption, this is effectively identical to the square of the time integral of the point spread function over the streak segment duration. The value of a is the integral over the sensor of this squared function. The expression for a is given in equation (3.14). Every term in the equation can be uniquely determined from the streak geometry. For a streaking object moving through the frame at a constant rate, the value of a is independent of the specific segment. Under the assumption of small pixels, the orientation of the streak within the image does not change the value of a. The x and y axes in equation (3.14) can then be selected to align x with the direction of motion of the streaking object and y perpendicular to that in the frame. The expression can then be simplified to equation (3.16) where l is the length of the streak in the frame in the same distance units as x and y.

$$a_i = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\int_{r_i}^{r_{i+1}} \frac{1}{2\pi\sigma^2}\exp\left(\frac{-1}{2}\begin{bmatrix} x-\mu_x(r) \\ y-\mu_y(r)\end{bmatrix}^T \sigma^{-2} I_{2\times 2}\begin{bmatrix} x-\mu_x(r) \\ y-\mu_y(r)\end{bmatrix}\right) dr\right]^2 dydx \tag{3.14}$$

$$a_i = \frac{1}{4\pi^2\sigma^4}\int_{-\infty}^{\infty}\exp\left(-\sigma^{-2}y^2\right)dy\int_{-\infty}^{\infty}\left\{\int_{r_i}^{r_{i+1}}\exp\left[\frac{-1}{2\sigma^2}(x-r)^2\right]dr\right\}^2 dx \tag{3.15}$$

$$a_i = \frac{1}{4\pi^{3/2}\sigma^3}\int_{-\infty}^{\infty}\left\{\int_{r_i}^{r_{i+1}}\exp\left[\frac{-1}{2\sigma^2}(x-r)^2\right]dr\right\}^2 dx \tag{3.16}$$

The inner integral over time in equation (3.16) can be solved analytically to yield the difference of two error functions, and this simplified form is shown in equation (3.17). Replacing the r_{i+1}

term with $r_i + \frac{l}{n}$, the expression can be simplified further. For any value of r_i, the integrand is horizontally shifted. Since the integral goes from negative to positive infinity, the choice of r_i is arbitrary. r_i can be chosen as zero, and a_i can be expressed as equation (3.18).

$$a_i = \frac{1}{8\pi^{1/2}\sigma} \int_{-\infty}^{\infty} \left[\text{erf}\left(\frac{r_{i+1} - x}{\sqrt{2}\sigma} \right) - \text{erf}\left(\frac{r_i - x}{\sqrt{2}\sigma} \right) \right]^2 dx \tag{3.17}$$

$$a_i = \frac{1}{8\pi^{1/2}\sigma} \int_{-\infty}^{\infty} \left[\text{erf}\left(\frac{\frac{l}{n} - x}{\sqrt{2}\sigma} \right) - \text{erf}\left(\frac{-x}{\sqrt{2}\sigma} \right) \right]^2 dx \tag{3.18}$$

The off-diagonal non-zero elements, b, can be found in a similar manner. They are the sum of the pixel-wise products of two adjacent streak segments. Once again, based on the assumed smeared Gaussian shape for the streak, all values of b should be identical to one another as the H matrix is independent of the individual streak segment intensities. Similar to the simplification used for the diagonal elements, the expression for b is first expressed as shown in equation (3.19). The innermost integrals can be solved with a similar approach as the one used for the diagonal elements, and the result is the product of the difference of two sets of error functions as expressed in equation 3.20. When these terms are multiplied out, this yields four error function products. Of these four, three have two different inputs for each function while one has identical inputs. Applying the assumption of long streak segments once again, the three unlike multiplied pairs sum together to a constant function of one independent of the value of x. The full expression then reduces a shown in equation (3.21). The x integral of this function over the full domain is simply $\frac{4\sigma}{\sqrt{\pi}}$, and the full equation simplifies to equation (3.22).

$$b_i = \frac{1}{4\pi^{3/2}\sigma^3} \int_{-\infty}^{\infty} \left\{ \int_{r_i}^{r_{i+1}} \exp\left[\frac{-1}{2\sigma^2} (x - r)^2 \right] dr \right\} \left\{ \int_{r_{i+1}}^{r_{i+2}} \exp\left[\frac{-1}{2\sigma^2} (x - r)^2 \right] dr \right\} dx \tag{3.19}$$

$$b_i = \frac{1}{8\pi^{1/2}\sigma} \int_{-\infty}^{\infty} \left[\text{erf}\left(\frac{r_{i+1} - x}{\sqrt{2}\sigma} \right) - \text{erf}\left(\frac{r_i - x}{\sqrt{2}\sigma} \right) \right] \left[\text{erf}\left(\frac{r_{i+2} - x}{\sqrt{2}\sigma} \right) - \text{erf}\left(\frac{r_{i+1} - x}{\sqrt{2}\sigma} \right) \right] dx \tag{3.20}$$

$$b_i = \frac{1}{8\pi^{1/2}\sigma} \int_{-\infty}^{\infty} \left[1 - \left(\text{erf}\, \frac{r_{i+1} - x}{\sqrt{2}\sigma} \right)^2 \right] dx \tag{3.21}$$

$$b_i = \frac{1}{2\pi} \tag{3.22}$$

With both of the terms a and b defined in terms of the streak's geometry, the recursive equation (3.13) can be simplified to the more computationally efficient summation form shown in equation (3.23). An induction based proof for the equivalence of these two representations is given below. To simplify the expressions, a is represented with c_1 and b^2 is represented with c_2.

$$f_n = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \binom{n-k}{k} c_1^{n-2k} c_2^k \tag{3.23}$$

$$\text{By definition } f_0 = f_{-1} = 0$$

$$f_1 = c_1 \qquad f_2 = c_1^2 - c_2$$

Now consider the case where both n and $n-1$ cases are assumed true. The summation forms for f_n and f_{n-1} can then be substituted into the recursive form for the f_{n+1} case.

$$f_{n+1} = c_1 f_n - c_2 f_{n-1} \tag{3.24}$$

$$f_{n+1} = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \binom{n-k}{k} c_1^{n-2k+1} c_2^k - \ldots$$

$$\sum_{k=0}^{\lfloor (n-1)/2 \rfloor} (-1)^k \binom{n-k-1}{k} c_1^{n-2k-1} c_2^{k+1} \tag{3.25}$$

$$f_{n+1} = c_1^{n+1} - \sum_{k=0}^{\lfloor n/2 \rfloor - 1} (-1)^k \binom{n-k-1}{k+1} c_1^{n-2k-1} c_2^{k+1} - \ldots$$

$$\sum_{k=0}^{\lfloor (n-1)/2 \rfloor} (-1)^k \binom{n-k-1}{k} c_1^{n-2k-1} c_2^{k+1} \tag{3.26}$$

$$\lfloor n/2 \rfloor = \begin{cases} \lfloor (n-1)/2 \rfloor & \text{if } n \text{ is odd} \\ \lfloor (n-1)/2 \rfloor + 1 & \text{if } n \text{ is even} \end{cases} \tag{3.27}$$

For the even n case, the summations can be combined as below.

$$f_{n+1} = c_1^{n+1} - \sum_{k=0}^{\lfloor n/2 \rfloor - 1} (-1)^k \left(\binom{n-k-1}{k+1} + \binom{n-k-1}{k} \right) c_1^{n-2k-1} c_2^{k+1} \tag{3.28}$$

$$f_{n+1} = c_1^{n+1} - \sum_{k=0}^{\lfloor n/2 \rfloor - 1} (-1)^k \binom{n-k}{k+1} c_1^{n-2k-1} c_2^{k+1} \tag{3.29}$$

$$f_{n+1} = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^k \binom{n-k+1}{k} c_1^{n-2k+1} c_2^{k} \tag{3.30}$$

$$\text{Since } n \text{ is even, } \lfloor (n+1)/2 \rfloor = \lfloor n/2 \rfloor \tag{3.31}$$

$$f_{n+1} = \sum_{k=0}^{\lfloor (n+1)/2 \rfloor} (-1)^k \binom{n-k+1}{k} c_1^{n-2k+1} c_2^{k} \tag{3.32}$$

For the odd n case, the summations can be combined as below.

$$\begin{aligned} f_{n+1} =& c_1^{n+1} - (-1)^{(n-1)/2} c_2^{(n+1)/2} - \ldots \\ & \sum_{k=0}^{\lfloor n/2 \rfloor - 1} (-1)^k \binom{n-k-1}{k+1} c_1^{n-2k-1} c_2^{k+1} - \ldots \\ & \sum_{k=0}^{\lfloor n/2 \rfloor - 1} (-1)^k \binom{n-k-1}{k} c_1^{n-2k-1} c_2^{k+1} \end{aligned} \tag{3.33}$$

$$\begin{aligned} f_{n+1} =& c_1^{n+1} + (-1)^{(n+1)/2} c_2^{(n+1)/2} - \ldots \\ & \sum_{k=0}^{\lfloor n/2 \rfloor - 1} (-1)^k \left(\binom{n-k-1}{k+1} + \binom{n-k-1}{k} \right) c_1^{n-2k-1} c_2^{k+1} \end{aligned} \tag{3.34}$$

$$f_{n+1} = c_1^{n+1} + (-1)^{(n+1)/2} c_2^{(n+1)/2} - \sum_{k=0}^{\lfloor n/2 \rfloor - 1} (-1)^k \binom{n-k}{k+1} c_1^{n-2k-1} c_2^{k+1} \tag{3.35}$$

$$f_{n+1} = \sum_{k=0}^{\lfloor n/2 \rfloor + 1} (-1)^k \binom{n-k+1}{k} c_1^{n-2k+1} c_2^{k} \tag{3.36}$$

$$\text{Since } n \text{ is odd, } \lfloor (n+1)/2 \rfloor = \lfloor n/2 \rfloor + 1 \tag{3.37}$$

$$f_{n+1} = \sum_{k=0}^{\lfloor (n+1)/2 \rfloor} (-1)^k \binom{n-k+1}{k} c_1^{n-2k+1} c_2^{k} \tag{3.38}$$

In both cases, the result is the summation result for the $n+1$ case. By induction, the given summation equation is the same as the recursive equation. With the formula, trial values of n can be tested and the value that returns the highest corresponding determinant value is the optimal number of states for fitting an observed streak to a stairstep model estimated streak.

Explicitly, the approach for finding the optimal value of n would be as follows. First, select

an initial small trial value of n and find the values of a_i and b_i with equations (3.18) and (3.22) respectively. This initial value can be one. Then compute the determinant with equation (3.23). Increase n by one and repeat the process recording each value of f_n along the way. The equation may not be unimodal, but there will be a point where the determinant approaches zero for larger n. At this point, the value of n that returned the greatest determinant is taken as the optimal value of n.

3.2.4 Computational Approach

The computational approach avoids the same assumptions required under the analytical approach by directly computing the integrals for each pixel and each streak segment. The most accurate and computationally costly approach to this method is to directly solve the integrals over each pixel for each feasible value of n. These results can then be used to build the H matrix for each trial value of n, and the determinant for each case can then be directly computed. A significant amount of computational effort can be saved by reducing the number of pixels the integration is performed for.

First, this can be done by only looking at the rectangular region of the image containing the streak and performing no integration on any points outside of this. Next, as the streak only occupies one of the image region's diagonals much of the rest of the region will integrate to near zero. Points beyond five point spread function standard deviations from the streak will have a negligible impact on the resulting model and can be safely assumed as zero. The lower the signal to noise ratio of the streak is, the tighter these bounds can be to reduce the number of integrated pixels.

While the previous approach does not impact the effective accuracy of the computation method, there are approaches that can significantly speed up the optimization that do have drawbacks depending on the particular scenario. The computation cost can be reduced by approximating the integration over each pixel. One approach is to approximate the incident intensity as locally linear over the area of the pixel. Instead of integrating over two spatial dimensions and time, just

integrating over time and finding the value at the center of each pixel allows linear regions to be constructed by defining surfaces based on three adjacent pixel centers. For square pixels, the distance between the closest pixel in the x and y directions will be the same, so the choice of which two neighbors to use to build the local linear model can be made arbitrarily with one on the x direction and one in the y direction. However, the choice should remain consistent between pixels so that no intensity values are over-weighted compared to the others. So long as the assumption that the streak is locally linear over the dimensions of a pixel holds, this approximation can save computational resources without losing significant precision. However, if the assumption does not such as when the pixel size is not sufficiently small relative to the point spread function, this linear approximation can fail to match the true geometry of the streak.

Algorithm 2 Algorithm for brute force calculation of the optimal number of streak segments.

$[i_0, j_0] \leftarrow$ Coordinates of the streak start.
$[i_1, j_1] \leftarrow$ Coordinates of the streak end.
$\sigma \leftarrow$ Point spread function standard deviation
$length \leftarrow \sqrt{(i_1 - i_0)^2 + (j_1 - j_0)^2}$
$maxI \leftarrow$ Number of rows in the image.
$maxJ \leftarrow$ Number of columns in the image.
$maxObs \leftarrow 0$, Maximum observability matrix determinant.
$optN \leftarrow 0$, The current best value for n
for $n = 1, \quad n < length, \quad n++$ **do**
    for $i = 0, \quad i < maxI, \quad i++$ **do**
        for $j = 0, \quad j < maxJ, \quad j++$ **do**
            for $k = 0, \quad k < n, \quad k++$ **do**
$$H[i + j \times maxI, k] \leftarrow \int_{t_k}^{t_{k+1}} \int_{j}^{j+1} \int_{i}^{i+1} \frac{1}{2\pi\sigma^2} \exp\left(\frac{-1}{2} \begin{bmatrix} x - \mu_x(t) \\ y - \mu_y(t) \end{bmatrix}^T \sigma^{-2} I_{2\times 2} \begin{bmatrix} x - \mu_x(t) \\ y - \mu_y(t) \end{bmatrix} \right) dydxdt$$
    $obsMatrix \leftarrow H^T H$
    if $\det(obsMatrix) > maxObs$ **then**
        $maxObs \leftarrow \det(obsMatrix)$
        $optN \leftarrow n$

3.2.5 Fitting the Data

Implementing either the analytical or computational method returns an estimated optimal number of states n. This is the number of steps in the stairstep function to represent all observable information in the streak without attempting to find estimates for more states than possible. With n known, the measurement matrix, H can be constructed from a modeled streak with the observed streak's measured start and end points and the point spread function's standard deviation. The full streak length is then broken down into n streak segments, and the predicted image of each streak segment is generated based on how a smeared Gaussian function would be seen on the image. Each streak segment's image is then rearranged into a column of the H matrix. The observed image is similarly rearranged into a column vector that is referred to with $\vec{y}$.

The estimated photometry of the streak is represented with the state vector, $\vec{x}$. Each value in this n element column vector is the intensity associated with its corresponding streak segment. This becomes the height of each step in the stairstep function. The values for the elements of $\vec{x}$ can be found via least squares by taking the pseudoinverse shown in equation (3.39). There the R matrix is the measurement uncertainty with each diagonal element representing the variance of a single pixel's value. The uncertainty for each of the n states can also be found using equation (3.1). This is the full covariance matrix for $\vec{x}$. These bounds can be further tightened by limiting their values such that no negative numbers are included in the feasible founds as pixel values cannot be negative.

$$\vec{x} = (H^T R^{-1} H)^{-1} H^T R \vec{y} \tag{3.39}$$

3.2.6 Application to Slitless Spectroscopy

Slitless spectroscopy is a tool primarily used in astronomy where a prism or diffraction grating and prism are placed in the optical system of a telescope to disperse the different frequencies of the received light. This results in each imaged point source appearing on the sensor as a streak.

The distribution of light along the streak depends on the frequency composition of the received light. Therefore, by extracting the intensity of light along the streak, one can determine the spectral composition of the light source. Current approaches for extracting this information do not take the information optimal approach [72]. The direct application of the proposed method for photometric light curve extraction to these streaks would provide an information-optimal approach to determining the spectroscopy to the best feasible resolution.

3.2.7 Limitations of the Model

The model makes several assumptions regarding the behavior of the imaging system. These assumptions are often true, but it is feasible that they may be violated under certain combinations of observation conditions and optical setups. Namely, these assumptions regard the bit-depth and saturation of pixels along the streak. The described method treats the value reported by each pixel as an accurate measure of flux the incident light on the pixel over the duration of the image integration. However, in real imaging sensors, the reported value is an integer between zero and two raised to the camera's bit depth minus one. This introduces a rounding error in the intensity values reported by each pixel. If the bit depth and source brightness are high enough, the reported intensity values will be high. The proportional deviation from the true intensity will be relatively small and well approximated by a continuous value. However, as the reported incident intensity decreases, the relative error imposed by this round off error grows.

The other major assumption is that the reported values do not saturate. The pixel cannot report a value greater than or equal to two raised to the power of the bit depth. This will cap the values reported by the sensor and the resulting streak will deviate in form from the true incident flux. Since the methods presented here assume that the incident flux is accurately reported, saturation can lead to an inaccurate estimation of the photometry. Specifically, the non-linear least squares regression along the streak length will report streak segment intensities lower than the true values. This is because the best fit curve for the streak with a saturated peak will fall below the best fit curve for a sensor without saturation at the same exposure.

3.3 Results

The proposed methods for estimating photometry are investigated in several ways to confirm they function as expected. First, simulated images are generated in order to verify that the proposed methods optimally extract the streak photometry in several scenarios. These results are also compared to non-optimal approaches to demonstrate what is missed when the approach is not information optimal. Next, the analytical and computational approaches are compared against one another to confirm that they provide similar results when the assumptions required for the analytical method are valid. Finally, the method is applied to extract the photometry of a imaged LEO.

3.3.1 Verifying Optimality

To confirm that the described method for extracting the photometry works as expected, the observed and estimated streaks can be compared. If the difference between the two does not contain any information about the streak, then the estimated streak represents all discernible information in the observed streak. To assess this, the appropriate number of streak segments for a streak with a known length and point spread function can be calculated using one of the methods discussed. The weighted least squares method shown in equation (3.9) is applied and the product of the H matrix and the calculated $\vec{x}$ is reshaped into an estimated streak. Then, a full estimate of the streak photometry based on the calculated number of streak segments can be used to build an artificial streak with the predicted photometry that shares the same endpoint coordinates and point spread function as the original observed streak. As both of these values are used in the calculation of the streak photometry, it is assumed that they are already known for the subject streak of interest. The difference between the original image and the estimated streak image can then be quantitatively assessed to insure that all information has been removed from the difference image. If the difference image is purely white Gaussian noise in both the pixel and frequency domains, no information remains in the image. This is because there is no information contained

in white Gaussian noise.

While the aforementioned test can ensure that the chosen number is streak segments is not too low to fully model the available information. It does not directly indicate if the chosen number is larger than necessary. Selecting a value too large can negatively impact the confidence in the values of the individual streak segment magnitudes as the method attempts to find estimates for unobservable states. One check to confirm that the selected value for the number of streak segments is not above the optimal value is to check the results for a model with one fewer streak segments. If this model successfully pulls out all available information in the streak, then either this model is more optimal than the larger value, or the streak segment magnitudes do not vary enough in the given streak to show the unmodeled information. By testing the method with simulated streaks, the photometry of the true streak can be controlled such that the second case is not a possibility. This ensures that reducing the number of streak segments appropriately tests for optimality.

3.3.2 Simulated Scenarios

Testing the two algorithms to find the optimal number of streak segments can be conveniently performed with simulated streaks. This involves choosing a point spread function that represents a realistic distribution and generating a large number of random values with that distribution. The number of photons generated at each time step is proportional to the intensity of the simulated source at that time. These values are then taken as coordinates on a image and binned corresponding to the pixel they would fall into. The mean of the distribution is then adjusted to model a smearing point spread function as the source slides across the image. At each time step, a new point spread function mean is found, and a new set of random photons are generated and added to a growing sum for each pixel. Additionally, to model the background noise, additional simulated photons can be generated with a uniform distribution across the image and count corresponding to the chosen background noise level to achieve a desired signal to noise ratio. This does not need to be performed at each time step, as identical results could be achieved by adding in the background noise at the end. Instead of generating many uniformly generated values, the result would

be better modeled by a uniform sum distribution for all the time steps. Time is continuous, so the resulting uniform sum distribution could be modeled most accurately by using normally distributed background noise counts for each pixel. To summarize the method, the algorithm for generating a simulated streak is given in algorithm 3.

Algorithm 3 Algorithm for generating simulated streaks.

$t_f \leftarrow$ Duration of the simulated image integration time.
$\delta t \leftarrow$ Time step size.$\delta t \ll t_f$
$t \leftarrow$ Current time step.
$PSF(\mu(t)) \leftarrow$ Source PSF with mean $\mu(t)$.
$P(t) \leftarrow$ Source photometric intensity at a given time.
$image \leftarrow$ 2D matrix values corresponds to a pixel count.
$noiseLevel \leftarrow$ The intensity of the background noise
$t \leftarrow 0$
while $t < t_f$ **do**
    for $k = 0, \quad k < \text{round}(P(t)), \quad k++$ **do**
        $[i, j] \sim PSF(\mu(t))$
        $[i_N, j_N] \leftarrow \text{Floor}([i, j])$
        $image[i_N, j_N] \leftarrow image[i_N, j_N] + 1$
    $t \leftarrow t + \delta t$
$noise \sim \mathcal{N}(0_{iMax \times jMax}, \text{diag}(noiseLevel_{iMax \times jMax}))$
$image \leftarrow image + noiseLevel \times noise$

When checking to see if the residuals between the original streak and the estimated streak leave any information in the image, there are two main parameters to check. One is the spatial distribution of residuals within the frame. Concentrated regions of either very high or very low residuals show up as bright or dark spots respectively. These exist because the estimated streak did not have enough states to model all the observable changes in the streak. The residuals for this image can be arranged in a histogram according to their value. If the distribution is Gaussian, the spatial residuals are distributed in a manner consistent with white noise. However, if information remains in the difference image, not all of the information may be spatially correlated to clearly show up in this plot. The information may instead be distributed periodically and this can be detected by taking the two dimensional Fourier transform of the difference image. Similar to the spatial residuals, both the real and imaginary parts of the Fourier transform should have a Gaussian distribution for their residuals if there is no information remaining in the frequency domain [68].

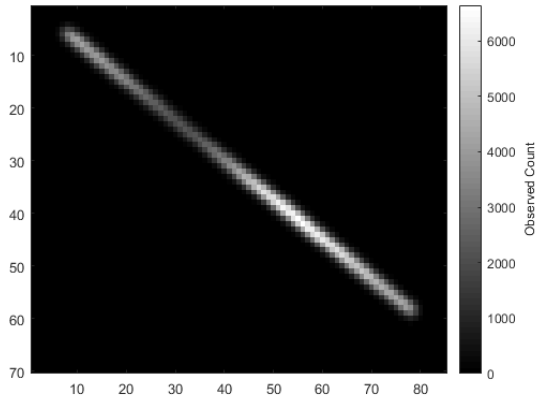

Figure 3.2a: High SNR artificial streak with sinusoidal photometry.

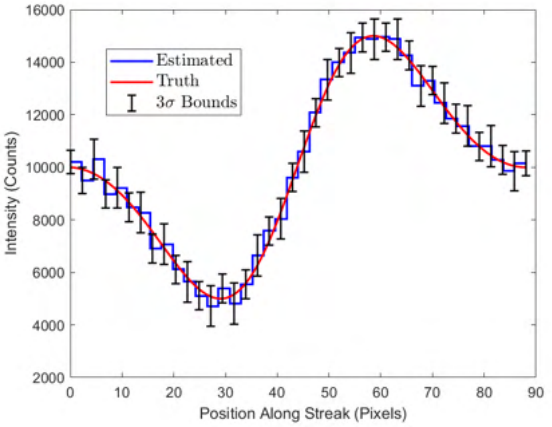

Figure 3.2b: Photometric light curve estimate of the high SNR simulated streak.

An artificial streak is generated using a known sinusoidal light curve and placed on a noisy background. The mean of the noisy background is subtracted from the image to simulate background subtraction pre-processing, so some pixels will have negative values. An example simulated streak with high signal to noise ratio is shown in figure 3.2a. For this streak, the optimal number of states calculated with the generally applicable calculation-based approach is thirty. Simulating a streak and using the weighted least squares method to find the photometric light curve yields a result that closely approximates the truth as shown in figure 3.2b. The difference between the original and estimated streak is shown in figure 3.2c. The frequency domain difference in both real and imaginary axes is shown in figures 3.2d and 3.2e respectively. These frequency domain plots are found by taking the fast Fourier transform of the difference image. The difference plots appear to be white Gaussian noise, to show this more explicitly, figures 3.2f through 3.2h show the histograms of values for the three difference plots. The histograms are fit with Gaussian curves, which accurately follows the distribution indicting that the distribution is truly Gaussian.

To show what happens when the number of states is too small, the same streak is estimated using an n value of five. Choosing such a low number allows the patterns that appear for a suboptimal choice of n to be clearly visible in all of the plots. The same issues remain detectable for closer to optimal values of n, but they require measures of the goodness of fit of the Gaussian to

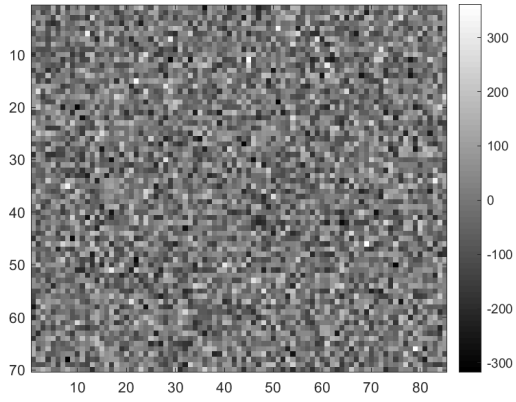

Figure 3.2c: Difference image between the high SNR simulated streak and the properly estimated streak.

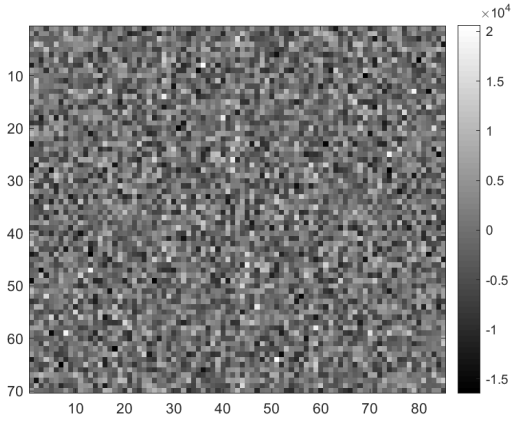

Figure 3.2d: Real component of 2D Fourier Transform of the high SNR difference image.

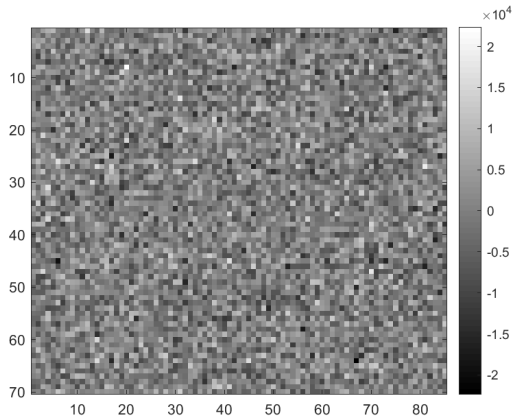

Figure 3.2e: Imaginary component of 2D Fourier Transform of the high SNR difference image.

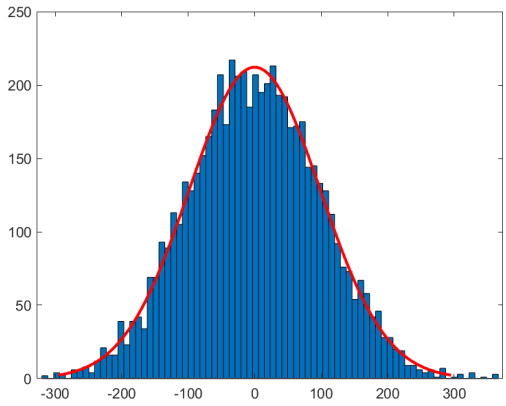

Figure 3.2f: Histogram and Gaussian fit of the high SNR difference image.

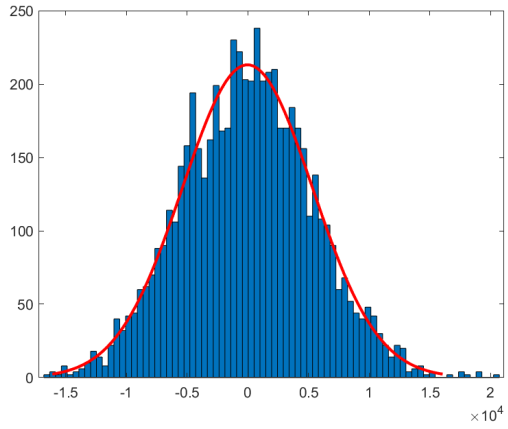

Figure 3.2g: Histogram and Gaussian fit of the real component of the Fourier transform of the high SNR difference image.

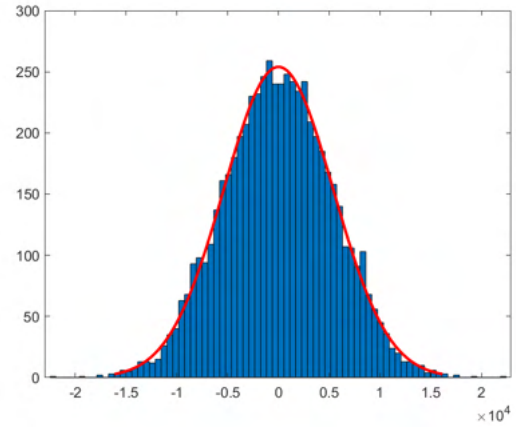

Figure 3.2h: Histogram and Gaussian fit of the imaginary component of the Fourier transform of the high SNR difference image.

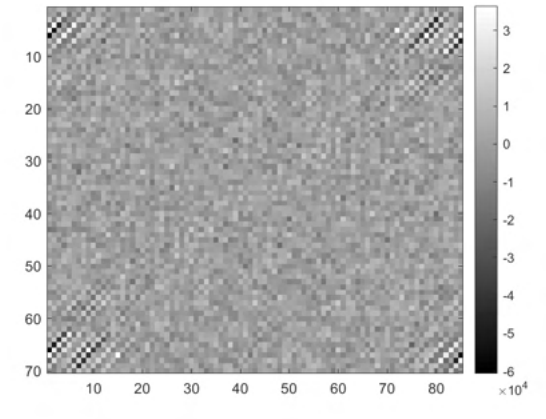

Figure 3.3a: Real component of 2D Fourier Transform of the suboptimal streak estimate and true streak difference.

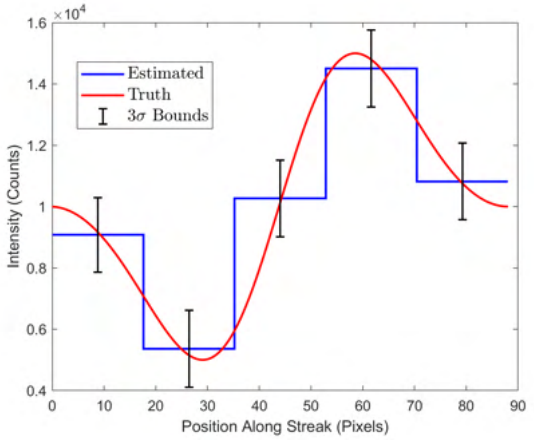

Figure 3.3b: Photometric light curve estimate with a suboptimal number of substreaks.

determine that they left information in the image. Figure 3.3a shows the real part of the frequency domain of the difference of the two streaks. There are clear structures in the corners of the image and the histograms further confirm the non-optimality as figures 3.3c and 3.3d show that the histograms of the differences are not well modeled by Gaussian curves.

A second streak similar to the first, but with lower signal to noise ratio ranging from two to six is also tested. This second streak is shown in figure 3.4a and the corresponding estimated photometric light curve and associated uncertainties are shown in figure 3.4b. The reduced signal to noise ratio significantly impacted the uncertainty of the estimates, but the estimates still fell within three standard deviations of the truth.

3.3.3 Method Comparison

The simplified analytical approach can be rapidly solved for a series of cases. Since the method only depends on the value of the point spread function standard deviation and the streak length, it is straight-forward to investigate the relationship between these two parameters as well as their overall response as their inputs change. Figure 3.5 shows the result returned by the simplified analytical approach for streaks of integer length between twenty and five-hundred and point spread

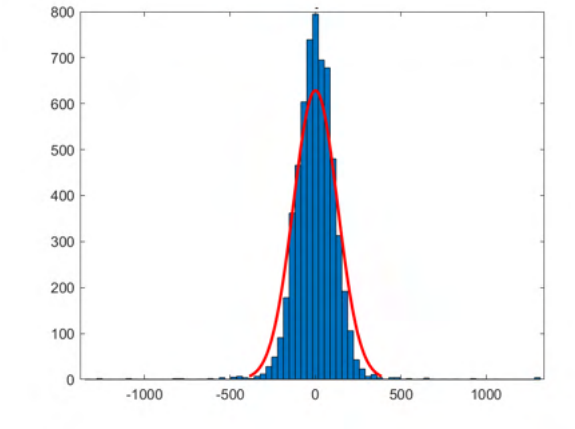

Figure 3.3c: Histogram and Gaussian fit of the suboptimal streak estimate and true streak difference.

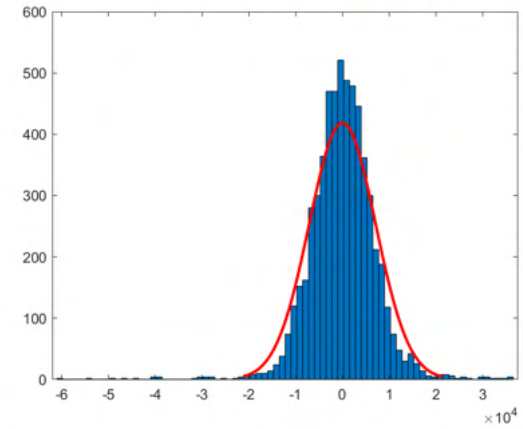

Figure 3.3d: Histogram and Gaussian fit of the real part of the Fourier transform of the poorly estimated streak.

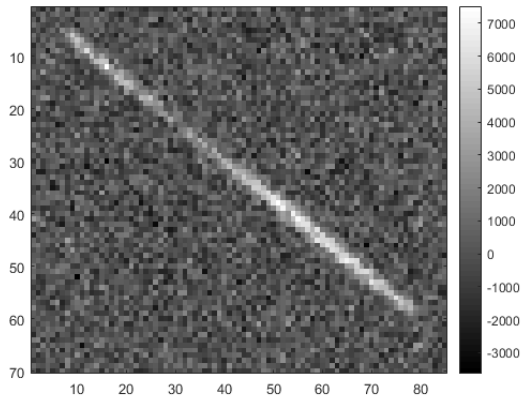

Figure 3.4a: Artificial streak with sinusoidal photometry and peak pixel SNR from 2 to 6.

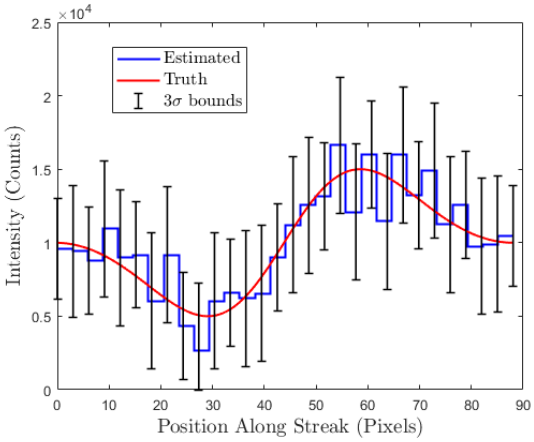

Figure 3.4b: Photometric light curve estimate of the 2 to 6 peak pixel SNR streak.

function standard deviation every half integer between one half and two. The plots show a linear relationship between the optimal number of streak segments and the length of the streak, and an inversely proportional relationship between the optimal number of streak segments and the standard deviation of the point spread function.

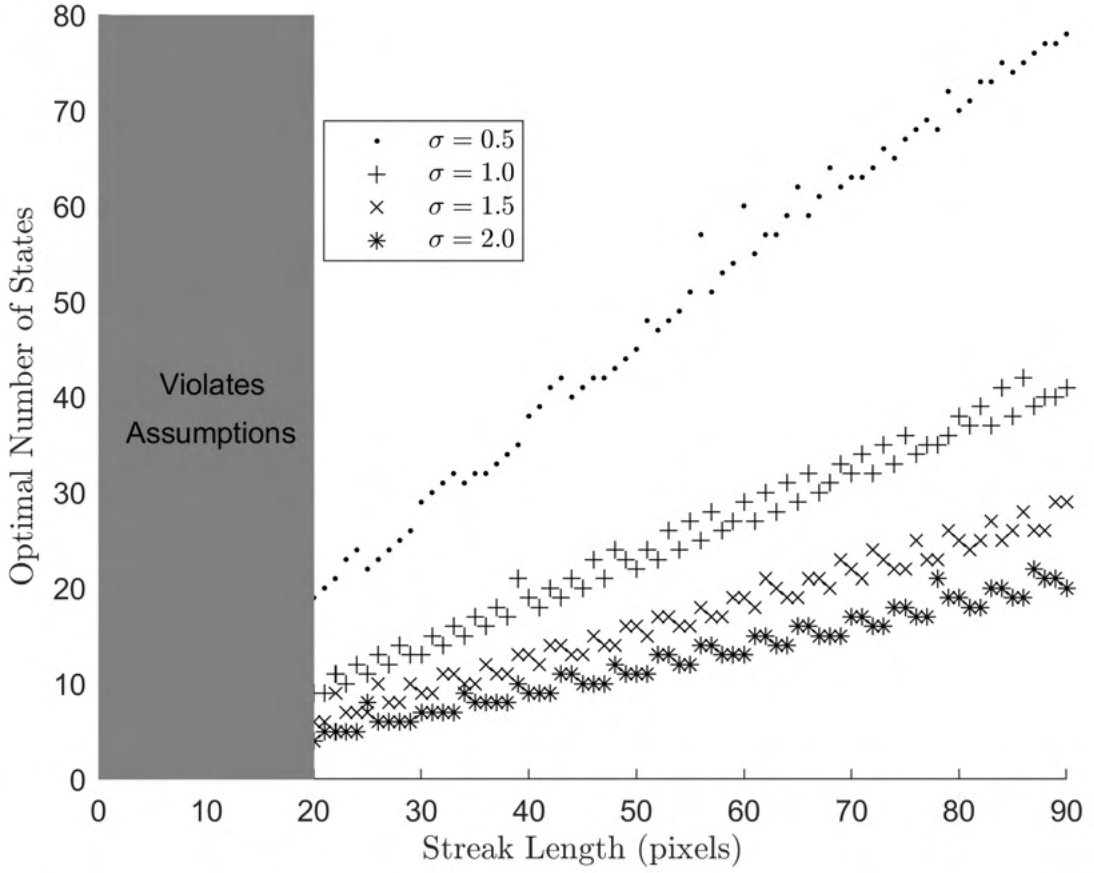

Figure 3.5: Simplified analytical results for the optimal number of states as a function of streak length and point spread function standard deviation.

Performing a simple linear regression on these four lines, the slopes are 0.16458, 0.08125, 0.05417, and 0.04167 for $\sigma = 0.5$, $\sigma = 1.0$, $\sigma = 1.5$, and $\sigma = 2.0$ respectively. These values are nearly inversely proportional to the standard deviations.

While the computational method applied in the previous section works for any case independent of the streak geometry, the analytical method that avoids the brute force approach does

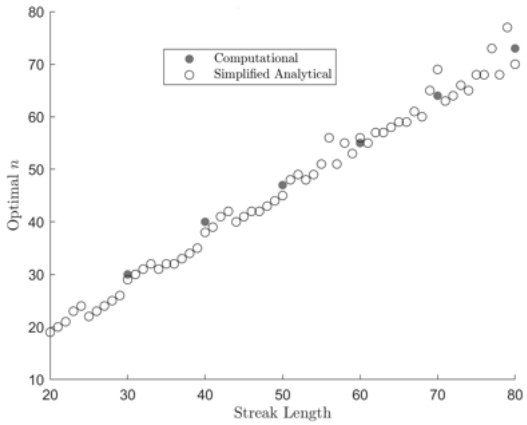

Figure 3.6a: Computational and simplified analytical methods for simulated streaks with 0.5 pixel standard deviation point spread functions.

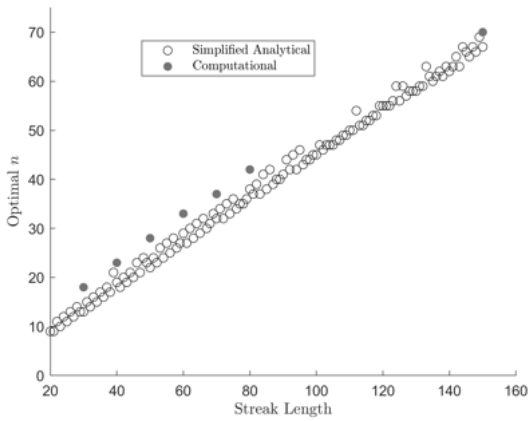

Figure 3.6b: Computational and simplified analytical methods for simulated streaks with 1.0 pixel standard deviation point spread functions.

not work for all cases. The analytical method makes heavy use of a set of assumptions outlined in the corresponding theory section. The simplified analytical approach requires that the streak is long and thin while remaining larger than the pixels themselves in both dimensions. However, it is not intuitively obvious where these assumptions break down in terms of streak geometry. To investigate this, simulated streaks are generated for a series of different lengths to check where the analytical approach diverges from the computational approach.

Figures 3.6a and 3.6b show the optimal number of streak segments found via the computational approach and the simplified analytical approach at a series of points. The standard deviation of the point spread function for all of these cases is set to 0.5 pixels in figure 3.6a and 1.0 pixel in figure 3.6b. The two methods provide similar results. However, in the case of the 1 pixel standard deviation point spread function, the optimal number of segments is under-counted by the simplified analytical approach until the streak length is sufficiently large. This matches the predictions for where the methods are effective. The simplified analytical approach assumes infinitesimal pixels and a long streak relative to the point spread function standard deviation. Accordingly, a larger difference between full computational results and the approximated results can be expected for streaks with larger point spread functions and shorter lengths.

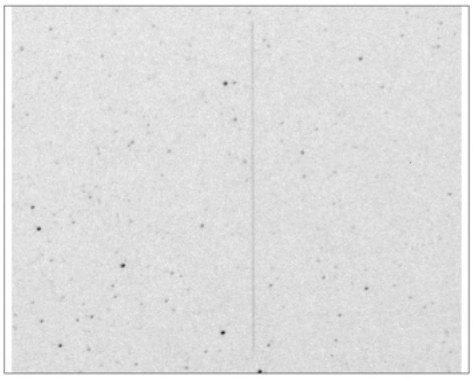

Figure 3.7a: Yaogan 3 as seen in Boulder, Colorado on November 8, 2019.

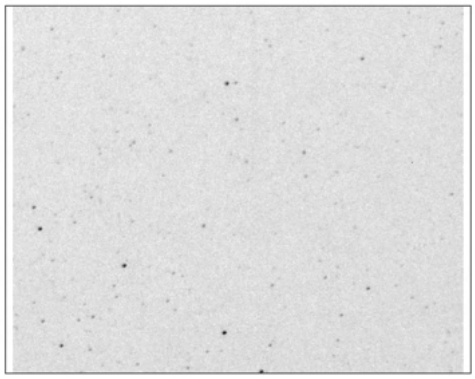

Figure 3.7b: Difference image of the original image and the estimated streak.

3.3.4 Results for Observed Streak

Finally, to confirm that the methods derived in this contribution can be applied in realistic situations, they are used on a real observed streaking object. The Chinese Yaogan 3 satellite, a low-Earth orbiting satellite in a nearly-polar orbit, was observed on the night of November 8, 2019 at 18:03 in Boulder, Colorado. The image was collected with a ZWO ASI 1800MM sensor with a Rokinon FE14M-E ultra-wide angle lens. The image integrated for ten seconds, and the resulting cropped streak left by the object is shown in figure 3.7a. The streak is 291 pixels long with a standard deviation of 0.1 pixels. The calculated optimal number of streak segments based on the simplified analytical approach is 89.

With the number of states known, the photometry of the streak can be extracted with a least squares regression. The extracted photometry and corresponding uncertainty are shown in figure 3.7d. The difference image of the observed and estimated streak is shown in figure 3.7b. Almost all of the original streak is removed in the difference image, and all that remains is some latent bloom in the surrounding pixels that is not accounted for by the implemented approach. The estimated streak itself is shown in figure 3.7c, and the intensity for all of these streak segments with their corresponding uncertainty bounds are shown in figure 3.7d.

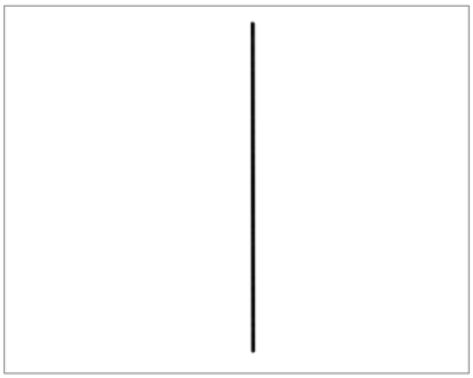

Figure 3.7c: Estimated Streak for Yaogan 3.

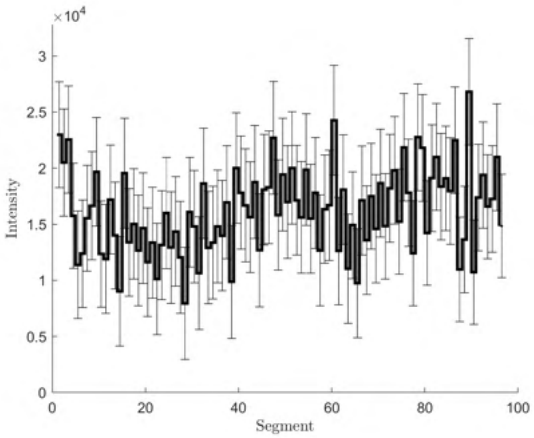

Figure 3.7d: Estimated photometry and 3 σ bounds for the light curve of the Yaogan 3 streak in figure 3.7a.

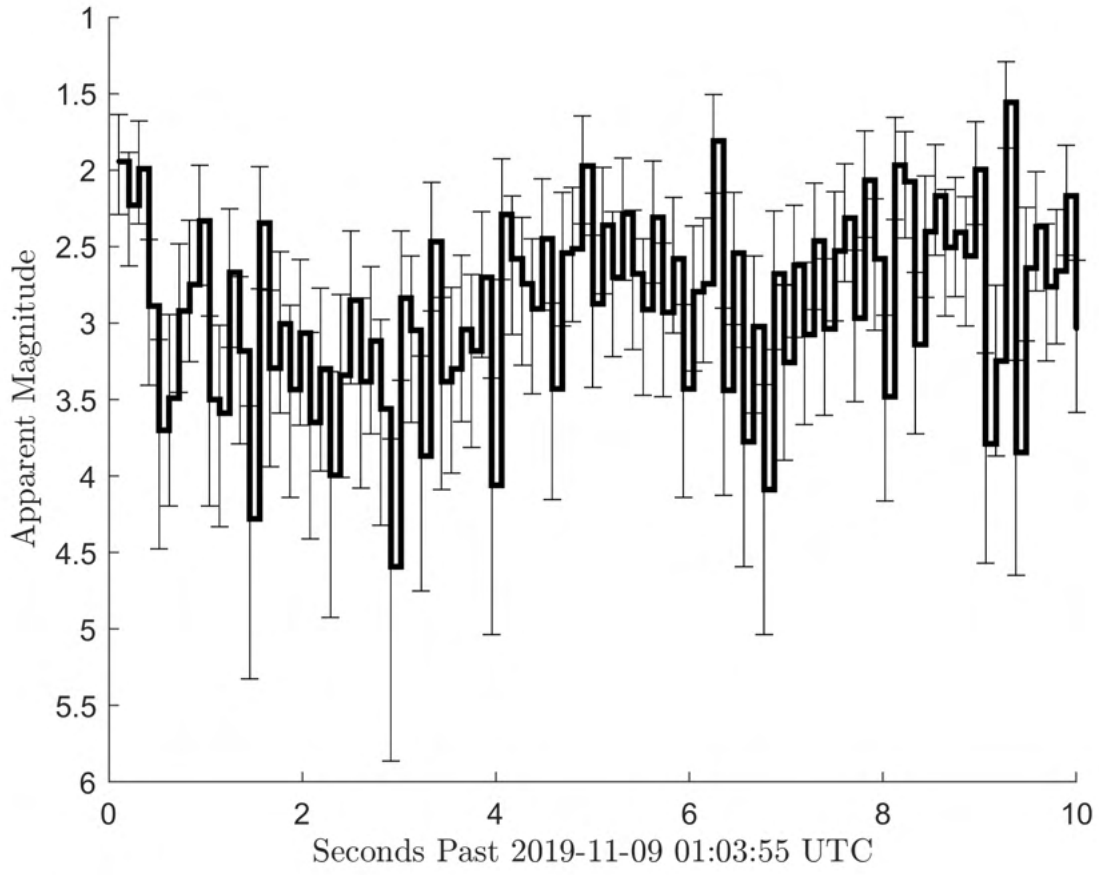

Figure 3.7e: Estimated apparent magnitude and 3 σ bounds for the light curve of the Yaogan 3 streak in figure. 3.7a.

3.4 Chapter Summary

An approach has been developed for optimally extracting the photometric information in a space object's imaged streak or a streak generated by slitless spectroscopy. Both a universally applicable computational approach as well as a computationally efficient, but simplified analytical approach have been developed. These methods can be used to find the information optimal number of segments to divide an observed streak into segments such that all available photometric information is accounted for, but the number of streak segments is not so high such that it attempts to estimate unobservable values. Performing a least squares regression on the model estimates the photometric light curve and provides values of the state uncertainty for each streak segment's

estimated intensity.

The two methods have been compared in order to determine the effective regimes for the simplified analytical approach as well as the behavior of the results as the length of the streak and standard deviation of the point spread function change. The simplified model accurately predicted the results of the fully computational model for long streaks. The simplified analytical approach was found to return values for the optimal number of streak segments directly proportional to the length of the streak and inversely proportional to the standard deviation of the point spread function.

After developing an understanding of the behavior of these methods, they were both applied to an observed streak produced by a low-Earth orbiting spacecraft. Both approaches were used to find an estimate of the optimal number of streak segments, and the results of the estimated photometry based on these subdivisions was reported and discussed.

Chapter 4

Novel IOD Approach with Admissible Regions

4.1 Initial Orbit Determination

There are a variety of initial orbit determination methods available. In terms of methods available for state estimation from optical data, there are two main types of algorithms. There are methods that accept only relative angle measurements between the observer and the target, and there are methods that accept the first time derivative of the heading as well. These are angles-only methods and angle and angle rates methods respectively [18]. In addition, many of these methods are more likely to converge in specific scenarios. Methods such as Laplace and Gauss tend to struggle with targets in low-Earth orbit as these methods perform best when the angle between observations is small. This narrows the feasible time windows between observations that allow these methods to remain useful [19]. Other angles-only methods like P.R. Escobal's double-r method work well with large angles between measurements. However, the iterative double-r algorithm has a relatively tight constraint on the initial guess for the target ranges that allows the method to converge to the correct solution [20]. The popular Gooding method often fails for space-based observers [19][21].

The variety in effective and ineffective regimes for these different approaches poses problems in itself. When little is known about a target object, it is not always clear which methods should be applied in order to find an accurate estimate of the true state. In addition, there are cases that do not perform particularly well under any initial orbit determination approach. One major case of this is for space-based observers measuring targets with large angles of separation between observations and little to no a-priori knowledge for generating initial guesses for the observation

ranges.

Most of the previously mentioned IOD algorithms are angles-only methods. These algorithms require a minimum of three observations to generate a state estimate as each measurement only provides two constraints while the system has six degrees of freedom. There are cases where a single optical observation of a space object can provide two additional constraints in the form of angle rate measurements. If the optical system translational and rotational rates of motion are known, the angle rates can be extracted from a single image. In this case, a minimum of two observations are required to estimate a full six degree of freedom state [73].

One promising approach for orbit determination is the L_n and similar J_n approaches developed by Rema Raymond Karimi and Daniele Mortari [30]. They showed in their 2011 paper how these angles-only algorithms could perform well in a variety of scenarios with only the zero vector as the initial range estimate.

The proposed method starts as an adaption of L_n rederived to accept angle rates and angles instead of only angles [30]. The L_n method uses a least squares optimization to find a solution that satisfies an equality based on the universal variable formulation of two body dynamics. The solution space contours of a cost function based on this approach show that the solution space has several local minima that can divert the algorithm from the correct solution for geometries where the target orbit lies inside the observer orbit. This cost function can then be searched with a global optimizing algorithm to find the solution corresponding to the true state.

The new proposed cost function has a unique non-trivial zero at the correct orbit solution. Because the system is overdetermined with the minimum two measurements, the global minimum will not approach zero if the two observations do not correspond to the same orbiting object. Applying a particle swarm descent method to the problem allows the system to converge to the correct state for any orbital regime. In its most natural implementation, this problem is a four-dimensional search. However, by projecting one measurement into the space of the other, it can be solved with a pair of two-dimensional searches which is much more computationally efficient and parallelizable. Applying admissible regions to both observations allows the state space to be

further constrained with a corresponding improvement to the computational efficiency. In addition, the particle swarm approach can report all of the feasible states in the multi-revolution case. The proposed method overcomes the shortcomings of many traditional initial orbit determination methods and has been empirically verified to find solutions to cases well-approximated by two-body dynamics. In addition, a Nelder-Mead optimization applied over the same cost function can show how the different optimization methods can impact computation time.

In section 4.2.2, the novel cost function is derived for measurements including both angles and angle rates. This cost function is modified with a penalty function to account for admissible regions. This penalty function is developed in section 4.2.3. Section 3.6 refines the cost function to modify the general underlying geometry of the surface and improve the performance of optimizers applied to the cost function. In terms of results, sections 4.3.1, 4.3.2, and 4.3.3 discuss the simulated performance of the algorithm for the single orbit case, multi-orbit case, and in response to measurement error respectively. Section 4.3.4 compares the results to other initial orbit determination algorithms. Finally, section 4.3.5 provides results for the algorithm's performance on observational data.

4.2 Theory

4.2.1 Initial Orbit Determination with Lagrangian Coefficients

Much of the work covered in this contribution is inspired by extending the work of Reza Raymond Karimi and Daniele Mortari in their 2010 paper [30]. Their approach to angles only initial orbit determination started with the Lagrange coefficients from the universal variable formulation of the two-body Kepler's problem. The approach applied the Lagrange coefficients to express the position state of the target spacecraft as a linear combination of the position state vectors at the previous and subsequent times. The position states then are broken down into their angular and range components. The full system can be expressed as a linear equation by pulling the ranges out as the independent vector. This is shown in equation 4.1. The c_i and d_i terms represent

the coefficients used to express the i^{th} position vector in terms of its previous and subsequent position states respectively. The $\hat{\rho}_i$ vectors are the unit vector headings from the observer position to the target position. The ρ_i terms are the ranges. Finally, $\vec{\xi}_i = \vec{R}_i - c_i\vec{R}_{i-1} - d_i\vec{R}_{i+1}$ where R_i corresponds to the observer position at the i^{th} measurement [30].

$$\begin{bmatrix} c_2\hat{\rho}_i & -\hat{\rho}_2 & d_2\hat{\rho}_3 & \vec{0} & \vec{0} & \dots & \vec{0} \\ \vec{0} & c_3\hat{\rho}_2 & -\hat{\rho}_3 & d_3\hat{\rho}_4 & \vec{0} & \dots & \vec{0} \\ \vec{0} & \vec{0} & c_3\hat{\rho}_3 & -\hat{\rho}_4 & d_4\hat{\rho}_5 & \dots & \vec{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vec{0} & \vec{0} & \vec{0} & \vec{0} & \vec{0} & \dots & d_{n-1}\hat{\rho}_n \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_n \end{bmatrix} = \begin{bmatrix} \vec{\xi}_2 \\ \vec{\xi}_2 \\ \vec{\xi}_3 \\ \vdots \\ \vec{\xi}_{n-1} \end{bmatrix} \tag{4.1}$$

4.2.2 Extension to Include Angle Rates

As the initial orbit determination method as presented by Karimi and Mortari accepts only relative angle of the target object as seen by the observer, the algorithm requires a minimum of three observations of the target object in order to provide an estimate of the target's orbit. The three observations each represent two degrees of freedom and the six total matches the full six degree-of-freedom translational state. However, some methods for observing space objects can readily yield angles and angle rates in a single measurement. For example, a non-tracking optical image taken with a relatively long integration time can show a streak instead of a point source for the target object's location. The direction and length of the streak can be used to estimate the target's angular rate with a single image. A measurement of both the target's angles and angle rates provides four degrees of freedom in a single measurement. In this case, only two measurements are required to establish an overdetermined estimate of the target's full six degree-of-freedom state. This both allows for a more efficient use of the collected data and reduces the minimum number of observations required to determine a state.

The angles-only method acts as a starting point for the derivation of a similar method that integrates angle rates as well. Both approaches start with two expressions for the position of the target object. One is defined with equation (4.2), while the other definition comes from equation (4.4). Instead of combining these equations for three different measurement times, two additional

equations for velocity are added in with the two position equations. These are an equation for velocity based on heading, heading rate, range, and range rate values as shown in equation (4.3). The other is the Lagrange coefficient equation for velocity as shown in equation (4.5). Equations (4.2) and (4.3) can be substituted into equations (4.4) and (4.5) for every instance of a target position or state vector. This will yield equations (4.6) and (4.7) which are only in terms of three types of parameters. These are the known parameters, $\vec{R}$ and $\dot{\vec{R}}$, the measured parameters, $\vec{p}$ and $\dot{\vec{p}}$, the range and range rate and the Lagrange coefficients that depend on the other parameters.

The Lagrange coefficients come from the universal variable formulation of the two-body Kepler problem. The variables f, g, $\dot{f}$, and $\dot{g}$ are derived directly from the dynamics of the physical system, and the formulaic definitions of these four terms are provided in equations (4.10) through (4.13) [74]. These four values are defined in terms of the translational inertial state vectors $\vec{r}_1$, $\vec{r}_2$, $\dot{\vec{r}}_1$, and $\dot{\vec{r}}_2$. These can be found from the known observer state, measured heading to the target and rate, and estimated range and range rate with equations (4.2) and (4.3).

$$\vec{r}_k = \vec{R}_k + \rho_k \vec{p}_k \tag{4.2}$$

$$\dot{\vec{r}}_k = \dot{\vec{R}}_k + \dot{\rho}_k \vec{p}_k + \rho_k \dot{\vec{p}}_k \tag{4.3}$$

$$\vec{r}_{k+1} = f_k \vec{r}_k + g_k \dot{\vec{r}}_k \tag{4.4}$$

$$\dot{\vec{r}}_{k+1} = \dot{f}_k \vec{r}_k + \dot{g}_k \dot{\vec{r}}_k \tag{4.5}$$

$$\vec{R}_{k+1} + \rho_{k+1}\vec{p}_{k+1} = f_k \left(\vec{R}_k + \rho_k \vec{p}_k \right) + g_k \left(\dot{\vec{R}}_k + \dot{\rho}_k \vec{p}_k + \rho_k \dot{\vec{p}}_k \right) \tag{4.6}$$

$$\dot{\vec{R}}_{k+1} + \dot{\rho}_{k+1}\vec{p}_{k+1} + \rho_{k+1}\dot{\vec{p}}_{k+1} = \dot{f}_k \left(\vec{R}_k + \rho_k \vec{p}_k \right) + \dot{g}_k \left(\dot{\vec{R}}_k + \dot{\rho}_k \vec{p}_k + \rho_k \dot{\vec{p}}_k \right) \tag{4.7}$$

$$p_{sr} = \frac{|\vec{r}_1 \times \dot{\vec{r}}_1|^2}{\mu} \tag{4.8}$$

$$\delta\nu = \arccos\left(\frac{\vec{r}_1 \cdot \vec{r}_2}{|\vec{r}_1||\vec{r}_2|} \right) \tag{4.9}$$

$$g = \frac{|\vec{r}_1||\vec{r}_2|\sin\delta\nu}{\sqrt{\mu p_{sr}}} \tag{4.10}$$

$$f = 1 - \frac{|\vec{r}_2|}{p_{sr}(1-\cos\delta\nu)} \tag{4.11}$$

$$\dot{g} = 1 - \frac{|\vec{r}_1|}{p_{rs}(1-\cos\delta\nu)} \tag{4.12}$$

$$\dot{f} = \frac{f\dot{g}-1}{g} \tag{4.13}$$

The terms containing either ρ or $\dot{\rho}$ can be collected onto one side of each of these equations as shown in equations (4.14) and (4.15).

$$\rho_{k+1}\vec{p}_{k+1} - f_k\rho_k\vec{p}_k - g_k\dot{\rho}_k\vec{p}_k - g_k\rho_k\dot{\vec{p}}_k = f_k\vec{R}_k + g_k\dot{\vec{R}}_k - \vec{R}_{k+1} \tag{4.14}$$

$$\dot{\rho}_{k+1}\vec{p}_{k+1} + \rho_{k+1}\dot{\vec{p}}_{k+1} - \dot{f}_k\rho_k\vec{p}_k - \dot{g}_k\dot{\rho}_k\vec{p}_k - \dot{g}_k\rho_k\dot{\vec{p}}_k = \dot{f}_k\vec{R}_k + \dot{g}_k\dot{\vec{R}}_k - \dot{\vec{R}}_{k+1} \tag{4.15}$$

The range and range rate terms can then be pulled out to reform these two equations into a single system of equations in a matrix form. This system of equations for the two observation case with both angles and angle rates is shown in equation (4.16). Note that the zero in the upper right corner of the matrix is represented as a vector since it represents a column of three zeros.

$$\begin{bmatrix} -f_1\vec{p}_1 - g_1\dot{\vec{p}}_1 & -g_1\vec{p}_1 & \vec{p}_2 & \vec{0} \\ -\dot{f}_1\vec{p}_1 - \dot{g}_1\dot{\vec{p}}_1 & -\dot{g}_1\vec{p}_1 & \dot{\vec{p}}_2 & \vec{p}_2 \end{bmatrix} \begin{bmatrix} \rho_1 \\ \dot{\rho}_1 \\ \rho_2 \\ \dot{\rho}_2 \end{bmatrix} = \begin{bmatrix} f_1\vec{R}_1 + g_1\dot{\vec{R}}_1 - \vec{R}_2 \\ \dot{f}_1\vec{R}_1 + \dot{g}_1\dot{\vec{R}}_1 - \dot{\vec{R}}_2 \end{bmatrix} \tag{4.16}$$

Just as the original least squares algorithm by Karimi and Mortari could be adapted to accept additional measurements beyond the three required, this angles and angle rates approach can be adapted to accepts more than two measurements. The relationship between the first and second measurement as shown in equation (4.16) defines the relationship between any pair of measurements. The same calculation can be performed with multiple observations beyond the initial two by extending the matrices in equation (4.16) with shifted forms identical to the relationship between the first and second measurements. The general form for a case with n measurements is shown in equation (4.17).

$$M = \begin{bmatrix} -f_1\vec{p}_1 - g_1\dot{\vec{p}}_1 & -g_1\vec{p}_1 & \vec{p}_2 & \vec{0} & \vec{0} & \vec{0} & \dots \\ -\dot{f}_1\vec{p}_1 - \dot{g}_1\dot{\vec{p}}_1 & -\dot{g}_1\vec{p}_1 & \dot{\vec{p}}_2 & \vec{p}_2 & \vec{0} & \vec{0} & \dots \\ \vec{0} & \vec{0} & -f_2\vec{p}_2 - g_2\dot{\vec{p}}_2 & -g_2\vec{p}_2 & \vec{p}_3 & \vec{0} & \dots \\ \vec{0} & \vec{0} & -\dot{f}_2\vec{p}_2 - \dot{g}_2\dot{\vec{p}}_2 & -\dot{g}_2\vec{p}_2 & \dot{\vec{p}}_3 & \vec{p}_3 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vec{0} & \dots & \vec{0} & -f_{n-1}\vec{p}_{n-1} - g_{n-1}\dot{\vec{p}}_{n-1} & -g_{n-1}\vec{p}_{n-1} & \vec{p}_n & \vec{0} \\ \vec{0} & \dots & \vec{0} & -\dot{f}_{n-1}\vec{p}_{n-1} - \dot{g}_{n-1}\dot{\vec{p}}_{n-1} & -\dot{g}_{n-1}\vec{p}_{n-1} & \dot{\vec{p}}_n & \vec{p}_n \end{bmatrix}$$

$$\vec{x} = \begin{bmatrix} \rho_1 \\ \dot{\rho}_1 \\ \rho_2 \\ \dot{\rho}_2 \\ \rho_3 \\ \dot{\rho}_3 \\ \vdots \\ \rho_n \\ \dot{\rho}_n \end{bmatrix} \qquad \vec{\xi} = \begin{bmatrix} f_1\vec{R}_1 + g_1\dot{\vec{R}}_1 - \vec{R}_2 \\ \dot{f}_1\vec{R}_1 + \dot{g}_1\dot{\vec{R}}_1 - \dot{\vec{R}}_2 \\ f_2\vec{R}_2 + g_2\dot{\vec{R}}_2 - \vec{R}_3 \\ \dot{f}_2\vec{R}_2 + \dot{g}_2\dot{\vec{R}}_2 - \dot{\vec{R}}_3 \\ \vdots \\ f_{n-1}\vec{R}_{n-1} + g_{n-1}\dot{\vec{R}}_{n-1} - \vec{R}_n \\ \dot{f}_{n-1}\vec{R}_{n-1} + \dot{g}_{n-1}\dot{\vec{R}}_{n-1} - \dot{\vec{R}}_n \end{bmatrix}$$

$$M\vec{x} = \vec{\xi} \tag{4.17}$$

Variables can be used to stand in for the matrix and vectors and the full system can be rewritten as a single linear equation of the form shown in equation (4.17) where M is the matrix, $\vec{x}$ is the vector of ranges and range rates, and $\vec{\xi}$ is the remaining vector. For a given series of measurements of the same space object, the least-squares solution for $\vec{x}$ can be found by taking the Moore-Penrose inverse of M and post-multiplying it by $\vec{\xi}$. These range and range rate values can be plugged back in to recalculate the Lagrange coefficients and therefore the elements of M and $\vec{\xi}$. In the results section of this chapter, it is shown that for an initial choice of zero for all ranges and range rates, this iterative least-squares converges to the correct solution in most cases. It is essential to remember that the initial assumption of zero only goes into finding the Lagrange coefficients for constructing the terms of M and $\vec{\xi}$. The first value of $\vec{x}$ is calculated from these terms and not initially assumed to be the zero vector.

The iterative least-squares method for estimating a full orbit state from angle and angle rate

measurements correctly finds the solution in many cases. However, convergence is not guaranteed, and there are cases where it fails such as those where the observer and target are both in or nearly in geostationary orbits. In order to approach the problem more rigorously, the least-squares method can be reformulated into a cost function. This will allow for new strategies for finding the target's state in addition to the iterative least-squares approach discussed previously.

4.2.3 Admissible Regions

Before discussing approaches to minimizing this cost function, there are additional factors that can be incorporated into the cost function to better constrain the feasible state space. Admissible regions allow the cost function to reflect a set of baseline restrictions on the derived orbit. These can be general restrictions for preventing non-physically realizable scenarios from being considered feasible states or they can be tighter bounds set by any a-priori knowledge about a target object's state. An example of the former would be a restriction on the periapse of the target object's orbit such that the target does not impact the Earth, while an example of the latter would a limit on the maximum observer to target range at the time of observation to reflect expected limits on the feasible detection distance between the two objects. A non-exhaustive list of potential admissible regions restrictions includes the orbit eccentricity, orbit energy, relative range, relative range rate, orbit periapse, and relative positioning of the target and observer such that illumination by the Sun would allow detection. Any or all of these restrictions can be calculated and applied as a penalty function to raise the cost of the associated states and reduce the size of the state space to search.

To prevent discontinuities in the solution space, the penalty function can be applied as an exponential to those states that violate one or more of the imposed restrictions. One practical way of implementing this is to find the amount by which a state violates a condition, η, and plug this into an exponential of the form shown in equation (4.18) with some multiplicative constant, c, and add this to the cost function. This multiplicative constant can be chosen for each admissible region restriction such that the costs for violating each restriction are balanced.

$$\exp(c\eta) - 1 \tag{4.18}$$

4.2.4 Particle Swarm Optimization

First, rather than making an assumption about the ranges and range rates to inform the elements of M and $\vec{\xi}$ and then finding $\vec{x}$ from this, the values for the ranges and range rates at each measurement are set to be the same in all of the terms. Equation (4.17) can be rewritten as a cost function by placing all of the terms on one side of the equation and recognizing that the difference of $M\vec{x}$ and $\vec{\xi}$ will only be the zero vector when a valid set of ranges and range rates are used to build all three variables. Accordingly, the difference of these terms will be non-zero for any solution that does not meet the constraints of the dynamical system. A single cost can be derived from this by taking the 2-norm of the difference. To balance the impact of inaccurate position and velocity states, a weighting matrix can be used to pre-multiply the difference vector before taking the norm. This weighting matrix is chosen to balance the impact of each component on the total cost. In Earth canonical units, the identity matrix is sufficient as the position and velocity components are similar in scale. In metric, choosing velocity weights of one thousand and positional weights of one can balance out the one thousand to one ratio between the scale of typical orbital radii and orbital velocities. The odd rows of the difference vector correspond to position errors while the even rows correspond to velocity errors. Weights can be chosen based on any prior knowledge of the target object's state. Alternatively, working in canonical units can account for much of the difference with unity weightings. The cost function with some diagonal weight matrix W is provided in equation (4.19). Independent of the choice for W, the cost is still zero if the solution is exact.

$$f = |W(M\vec{x} - \vec{\xi})| \tag{4.19}$$

With the problem rewritten as a cost function, a number of new approaches open up. One of the approaches is to perform a particle swarm optimization over the space of possible ranges and range rates. A particle swarm optimization starts with some number of randomly or strategically

placed particles spread out through the solution space of a cost function. Each step of the algorithm has the particles move in a random direction with probabilities weighted by both the minimum value the particle has encountered on the cost function as well as the minimum value found by any particle in the swarm. As the algorithm continues, particles are drawn to the global minimum so long as at least one particle has begun to explore the region around the global minimum. As more particles are pulled into this well, they are able to explore it more thoroughly with each iteration. There are a number of tunable parameters that define both how quickly the algorithm converges and how likely it is to fail and miss the global minimum.

The most direct implementation would be to perform a particle swarm optimization over a space with dimension equal to the number of elements in $\vec{x}$ in order to have a term for the range and range rate for each measurement. In its most basic form, the bounds on the particle swarm could be set to safe feasible limits with a maximum range of perhaps twice the Earth's Hill sphere and range rate of several kilometers per second to match.

However, admissible regions can be applied to give much tighter constraints on both of these limits. Applying admissible regions requires some basic assumptions about the target object's state, but they significantly reduce the feasible solution space. For tracking expected Earth-orbiting satellites, the target can be assumed to be in an elliptical orbit that does not impact the Earth. This provides constraints in the object's semi-major axis, orbital eccentricity, and radius of periapse.

.

Another tool for significantly constraining the solution space is to only find the range and range rate for a single observation and project this state to the other observation times to find their corresponding ranges and range rates from this assumption. This reduces the dimension of the solution space to two with one dimension corresponding to range and the other range rate for the first observation. This is possible since the measured heading and heading rate with a selected range and range rate fully defines the state at a time, and future states can be estimated by propagating the orbit forward in time by solving Kepler's equation. While there is no analytical solution to Kepler's equation (4.20), applying a root-finding algorithm due to Laguerre can find

the solution to machine precision with only a few iterations [75].

$$M = E - e \sin E \tag{4.20}$$

When an erroneous estimated state is projected to future times, the expected measurement of the projected state may not match the measured heading and heading rate. The simplest implementation of the cost function can ignore this error since erroneous states will be penalized in equation (4.19). However, it is also reasonable to find the angular distance between the observed and expected measurement and incorporate this difference into a new cost function that penalizes both error in equation (4.19) and error in the expected measurement. The contours of the solution space can be manipulated by choosing weights for each of these terms. An alternative cost function of this form is provided in equation (4.21). Errors for both the forward and backward projection of a chosen range and range rate are included for each measurement and the user-defined weightings are labeled as w_i. The predicted headings are denoted as $\hat{p}_i$.

$$f = \sum_{i=2}^{n} w_i \arccos\left(\hat{p}_i \cdot \vec{p}_i\right) + |W(M\vec{x} - \vec{\xi})| \tag{4.21}$$

Implementing a particle swarm optimization over the higher dimensional solution space would have been prohibitively costly in terms of computational resources. The number of particles required to search the solution space grows exponentially with the number of dimensions in the solution space [76]. With this reduced two-dimensional solution space, a particle swarm optimization can be run with much greater efficiency. Furthermore, the admissible region constraints can provide fairly tight bounds on this two-dimensional solution space depending on how much information the designer is able to assume. By using a universal variable formulation for both the Kepler's problem solver and in the derivation of the cost function itself, the approach does not need to assume a closed elliptical orbit. Rather, this method can be used to perform initial orbit determination on orbits corresponding to any conic section.

4.2.5 Nelder-Mead Optimization

With the cost function approach to initial orbit determination discussed in the previous section, any derivative-free optimization method can be applied to solve the problem. One particularly attractive alternative to the particle swarm optimization is the Nelder-Mead simplex-based approach [77]. In the case of a two-dimensional state space, the Nelder-Mead algorithm tracks the values reported by three points as they explore the cost function space. The simplex formed by these three points is manipulated by repositioning the three vertices through a method that brings the simplex as close as possible to the a minimum.

The method stats with the initial placement of the three vertices in the state space. A-priori knowledge regarding the likely state of the target can be used to choose an intelligent start point. In this case, the vertices should surround the minimum to maximize the likelihood that the optimizer will converge [78]. With no a-priori knowledge, the vertices can be chosen randomly. The implemented method assumes nothing and selects random states for each vertex. The cost for each vertex is checked, and vertices outside of the bounds of the penalty function are re-selected until a point inside the bounds of the penalty function is found.

In each iteration of the algorithm, the costs at each of the vertices and a point projected outside of the simplex and inside the simplex are compared. The vertex with the highest cost is moved to one of these new points, and the process is repeated until the simplex converges down to a minimum [79]. In order to ensure that the algorithm converges to the correct minimum, the full process is repeated several times. This is performed a maximum of one hundred times or when the algorithm exits when it finds a point with sufficiently low cost or after it hits a maximum number of iterations and returns the state with the lowest cost on any iteration. Repeating the full algorithm allows for multiple opportunities to generate an appropriate set of initial conditions so that the encountered minimum is the unique global minimum that corresponds to the solution. The choice of one hundred as a maximum is adjustable depending on the goals and computational limitations of the implementation.

4.2.6 Solution Space

Restructuring the initial orbit determination algorithm as a cost function opens the question of what form the solution space takes. This includes the number of minima that an optimization method might find in addition to the true solution. These additional convex regions of the cost function could serve to draw away an optimization algorithm from the true solution and have it report erroneous results. Similarly, there is the question of how deep these minima go. The cost function is designed to return zero for a solution that perfectly matches the observed measurements and dynamics. If the other minima are significantly above zero, this could be integrated into the optimization algorithm so that it can leave these regions if the local minimum does not approach a sufficiently small value. Beyond these two quantifiable measures of the solution space, the contours of the cost function can show if in some cases the minimum is not unique in the case of degenerate observations.

Figures 4.1 through 4.6 show the contours generated by the the two cost functions for a variety of scenarios. The first cost function is referred to as the projection cost function. This cost is based on taking the first measurement and an a range and range rate guess. This forms a full state estimate that is projected forward in time to the time of the second measurement. The cost is then taken as the angular error between the angle measurement at this second time and the expected measurement based on the range and range rate guess. The process is repeated with the second measurement projected back in time to the first measurement. These errors are added together to give the estimate's cost. The second cost function is the L_n+Rates cost function developed in the Theory section. A penalty function based on admissible regions is added to both of these cost functions. The admissible regions penalty function takes in the state estimate and checks that the range, range rate, radius of periapse, semi-major axis, and eccentricity are valid based on what is known about the system. For these figures, the bounds are set to require that the target is orbiting the Earth without impacting it and the eccentricity is below 0.2. To ensure a smooth transition between penalized regions and non-penalized regions, the penalty function takes

an exponential form.

When a state estimate falls outside the bounds defined by the penalty function, the difference between the state and the bound is multiplied by a weight to increase the penalty associated with erroneous eccentricities and velocities and balance them with the penalty for out of bounds ranges and semi-major axes. For Earth-orbiting satellites in metric units, distances tend to be about three orders of magnitude larger than velocities and velocities tend to be at least an order of magnitude greater than eccentricity. This acts as a guide for choosing weights so as not to allow errors of one type dominate the others. For example, if calculations are performed in units of kilometers and kilometers per second, a position in the inadmissable region by one kilometer is a minor infraction compared to a velocity penalty of one kilometer per second. Given that position values tend to be three orders of magnitudes greater than velocity values in low Earth-orbiting cases, it would be reasonable to weight the velocity error in these units by a factor one thousand and the position error by one. If the targets are expected to be in geosynchronous orbits, a ten-thousand to one ratio would be more fitting as that follows the relative ratio of the expected values. The penalty is then taken as the exponential of the sum of these weighted errors. One is then subtracted from the result so that the penalty is zero when all terms are within their respective bounds. The exponential form ensures that estimates that are far out of bounds are appropriately penalized.

4.2.6.1 Non-Coplanar Case

The four investigated cases correspond to all combinations of geosynchronous and low-Earth orbiting observers and targets in the single-revolution case. A red dot is placed on the image where the true solution is located. The LEO orbit for both the target and the observer is based on a nearly circular orbit near-equatorial orbit. Specifically, the orbit is at an altitude of 400 kilometers and the inclination is zero for the observer and 10 degrees for target. The geosynchronous orbit is defined as one with a semi-major axis of 42164 kilometers and a identical approach of zero degrees for the observer and 10 degrees for the target object. This prevents the case from being coplanar. The coplanar case is developed in the following section.

The penalty bounds determine the overall shape of the investigated region. In the case where both the observer and target are in a low-Earth orbit, the feasible states are confined to a crescent in range and range-rate space. A valley of low cost runs through this crescent. However, for the non-coplanar case, the cost only reaches zero for a unique non-trivial point that corresponds to the correct range and range-rate. The responses for both cost functions are similar, however, the exact line traced out by this valley is shifted for each case. In the projection error case, there are two distinct curves that form the cost valleys. In the L_n+rates function, there is a single valley that winds through the feasible region.

For the case where both the observer and target are confined to geosynchronous orbits, the crescent shape remains, but a small circular region with negative range-rate opens up with no penalty. However, this local minimum remains well above the global minimum still confined to the crescent region. Once again, this structure is similar for both cost functions. One detail that differs is the cost function response near the trivial solution. This is zero in both cases. However, in the projection case, this point around rise faster than in the L_n+Rates case.

The two remaining cases are for observers and targets in different orbits. In both of these cases, three distinct local minima show up. One is about the trivial solution, while the other two are non-trivial. Similar to the double geosynchronous case, one of these is a local minimum that does not fall to the global minimum found only at the correct solution and the trivial solution. Figures 4.5 and 4.6 show the LEO to GEO scenario. The GEO to LEO scenario is not shown. However, it is structurally similar to the LEO to GEO scenario. The main differences are small changes in the shape to the penalty function bounds and the location of the solution within the solution space.

4.2.6.2 Coplanar Case

The inclination of the second object is zeroed out for the LEO observer and LEO target case. While this does not significantly shift the over-all shape of the permissible region as defined by the admissible region penalty function, this does change the geometry of the contours inside

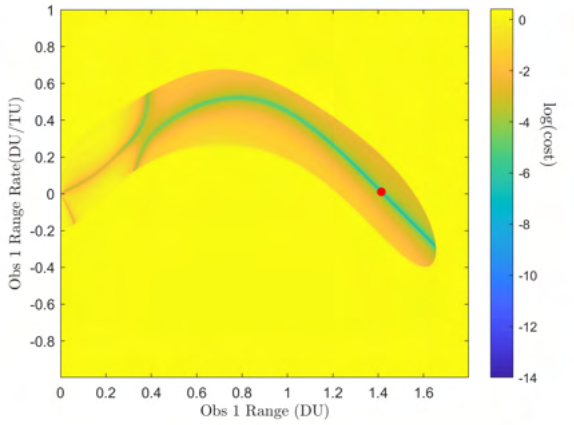

Figure 4.1: LEO observer to LEO target measurement error log(cost) contours.

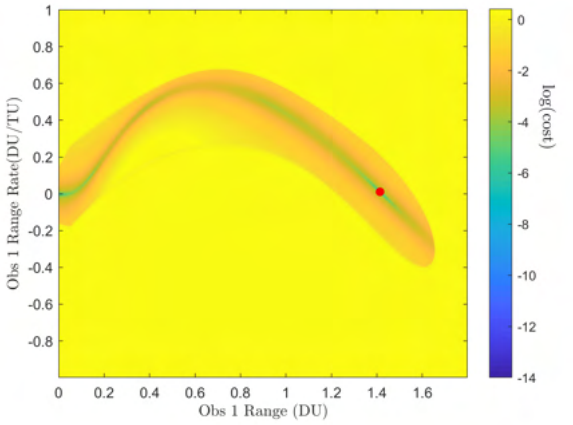

Figure 4.2: LEO observer to LEO target L_n+Rates log(cost) contours.

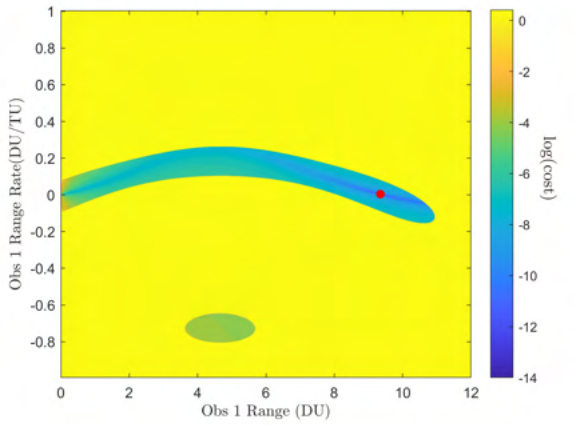

Figure 4.3: GEO observer to GEO target measurement error log(cost) contours.

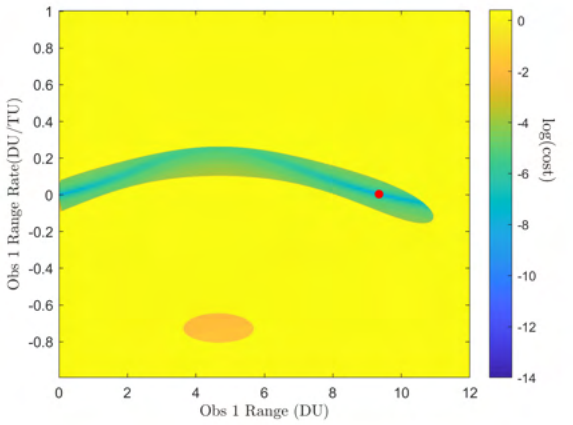

Figure 4.4: GEO observer to GEO target L_n+Rates log(cost) contours.

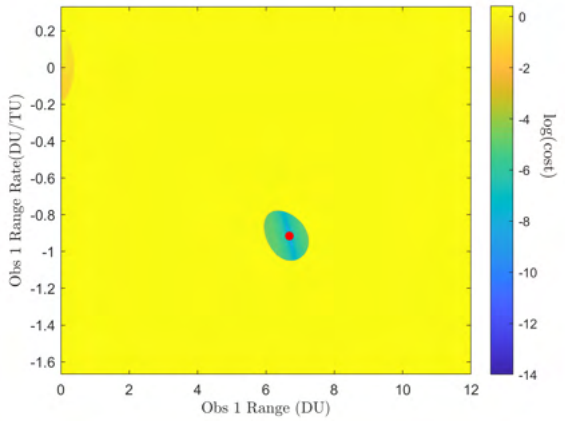

Figure 4.5: LEO observer to GEO target measurement error log(cost) contours.

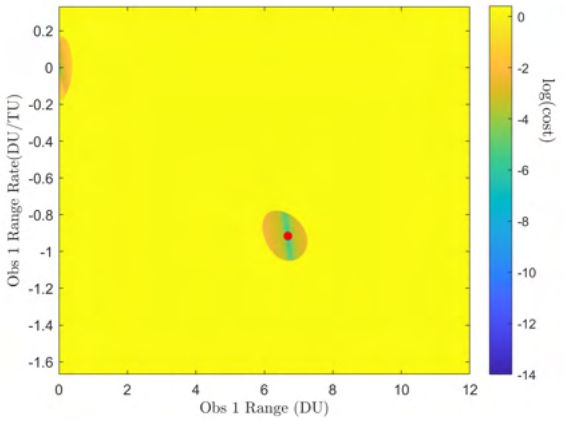

Figure 4.6: LEO observer to GEO target L_n+Rates log(cost) contours.

the permissible region. The projection-based cost function shows two curves of zeroes in the cost function, and as a result, the global minimum is no longer the unique solution. However, the L_n+Rates cost function does not face this same issue. In the coplanar case, the L_n+Rates cost crescent contains a unique global minimum and two additional non-trivial local minima. Since both cost functions correctly return zero for the correct solution, they can be combined together to generate a new cost function with the benefits of both geometries. The result for this combined cost function for the same LEO to LEO case is shown in figure 4.9. This function keeps the unique global minimum in the coplanar case from the L_n+Rates cost function and the favorable geometry around the trivial zero from the projection cost function. This favorable geometry keeps the cost around the trivial solution higher to help push the optimization function away. The trivial solution can be manually excluded using the admissible regions limits. However, the projection cost function helps keep this bound less attractive for the optimization algorithm.

The geometry for this combined cost function is favorable in many cases, but it is not guaranteed as the relative positions of the projection and L_n+rates cost function minima shift for different orbit scenarios. In some cases, the combined cost function contours remains more similar to that of the target measurement error alone contours. In these cases, while there may still remain an objective minimum, the long and shallow valley that contains it may also have other local minima that confound both the particle swarm and Nelder-Mead approaches as they descend. For one thousand Monte-Carlo simulated cases of coplanar LEO to LEO observations, 58.4 percent of the cases converged to the correct state. These were generated with the values in table 4.1.

Similar to the LEO observer and LEO target case, the Geostationary observer and target case does not have a unique solution for either the measurement projection or L_n+rates cost function. The curves are similar in structure to those for the LEO case. However, the L_n+rates case has shifted from several discrete minima to a continuous curve of minima as shown in figure 4.11. When the two cost functions are summed, the unique non-trivial intersection of these curves of minima align with the true solution to the problem. In the geostationary case as well, the combined cost function can provide the unique correct answer to the orbit determination problem. This

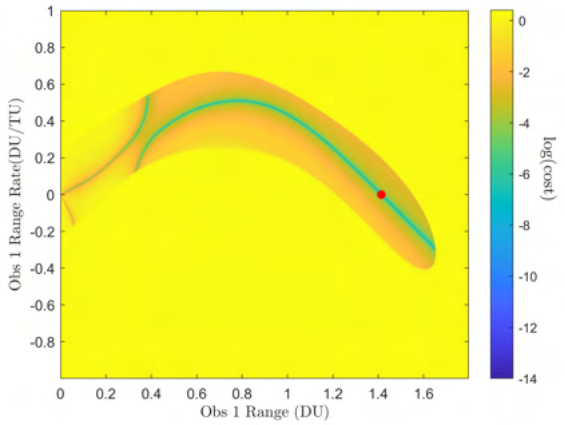

Figure 4.7: Coplanar LEO observer to LEO target measurement error log(cost) contours.

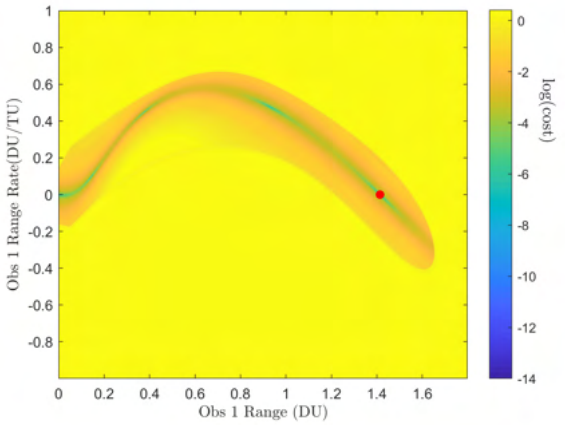

Figure 4.8: Coplanar LEO observer to LEO target L_n+Rates log(cost) contours.

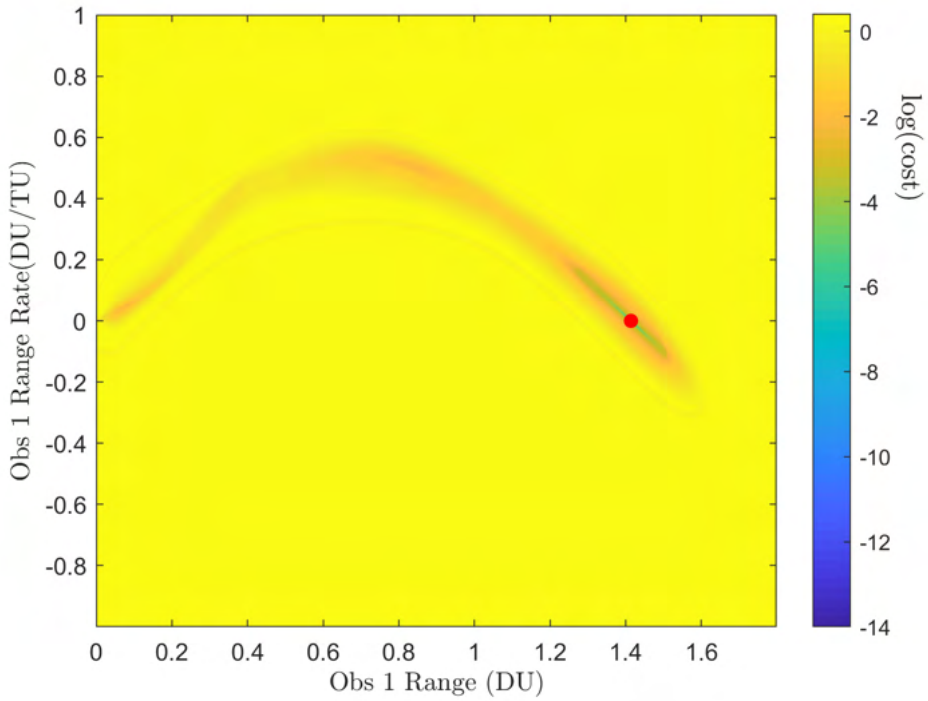

Figure 4.9: Coplanar LEO observer to LEO target log sum of the two cost functions.

occasionally suffers from the same issue of poor cost function geometry. For one-thousand Monte-Carlo simulated cases with the cases based on the values from table 4.1, 84.7 percent converged to the correct solution.

4.3 Results

4.3.1 Simulated Test Cases

A Monte Carlo method is applied to test the developed initial orbit determination methods under a variety of scenarios. Both the particle swarm approach as well as the Nelder-Mead algorithm are tested. Each one is run with the same set of one thousand observer and target pairs by generating the cases with a seeded random number generator. The same four scenarios as investigated in the cost function contour study are used again. That is each possible combination of low-Earth and geosynchronous orbits to see how each domain responds to both algorithms.

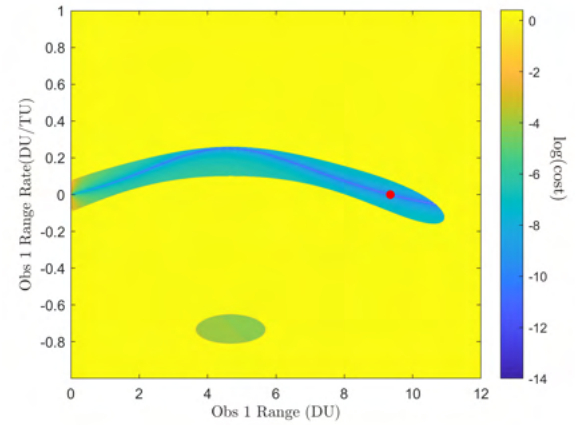

Figure 4.10: Coplanar GEO to GEO measurement error log(cost) contours.

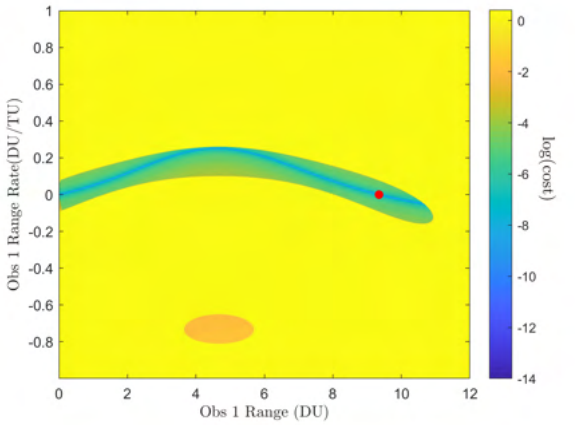

Figure 4.11: Coplanar GEO to GEO L_n+Rates log(cost) contours.

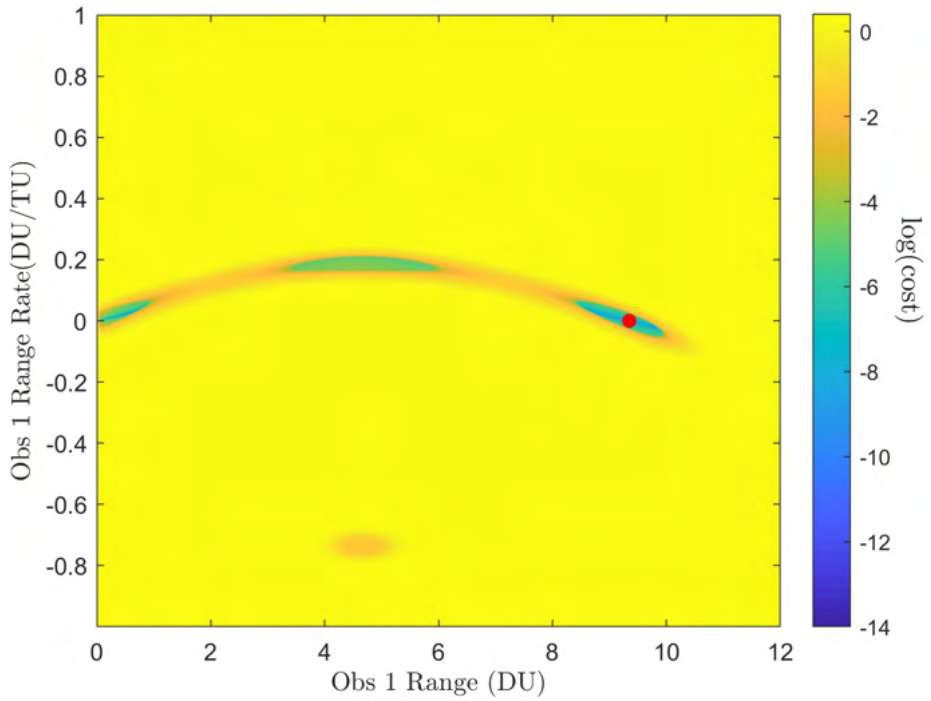

Figure 4.12: Coplanar GEO observer to GEO target log sum of the two cost functions.

The Monte-Carlo randomly generated observer and target orbits using the values from the table below. The values with means and standard deviations are generated from a normal distribution. The terms with a minimum and maximum are generated using a uniform distribution. The results of all the cases are plotted in a series of scatter plots. Figure 4.13 shows the results for the particle swarm applied to the LEO observer and LEO target and figure 4.14 for the Nelder-Mead algorithm.

The x-axis for the scatter plots is the time in Earth canonical time units (TU) between the first and second observation. One TU is equivalent to 806.8042 seconds. This comes from the mean motion of a theoretical orbit at the surface of the Earth. The y-axis results on each scatter plot is the time in seconds required to find the solution for each case. In both algorithms, this value is highly dependent on chosen parameters that balance the algorithm's ability to converge to the correct solution with the time required to compute the result. Finally, the color and shape of the point on the scatter plot corresponds to whether the estimated state matched the true state to

Type	a Mean	a STD	e Max	i Mean	i STD	$\Omega/\omega/\nu$ Min	$\Omega/\omega/\nu$ Max	Δt Min	Δt Max
LEO	8413 kilometers	100 kilometers	0.001	0°	45°	0°	360°	0 s	3600 s
GEO	42164 kilometers	100 kilometers	0.001	0°	7°	0°	360°	0 s	3600 s

Table 4.1: Monte-Carlo parameters for randomly generating orbits.

within one percent in all six translational state components. The gray squares had at least one state component that did not match the true state to within one percent. The blue circles represent estimated states that correctly match the true state on all components to one percent.

Both the particle swarm and Nelder-Mead methods converge toward a solution as they iterate. However, in a practical implementation there will be measurement error in the angle and angle rate data that the algorithm ingests as well as potential correlated error. Assuming the measurement error is well-approximated by a Gaussian, this can all be represented in a four by four measurement covariance matrix. This makes it impractical to define a convergence criterion based on some minimum value for the cost function. Due to this, the global minimum of the solution space may never reach a hard-coded stopping point and the algorithm would not be able to adjust to different levels of certainty in the input measurements. To account for this, the convergence criteria can be defined with the Mahalanobis distance. Either algorithm can accept a measurement covariance matrix input. The Mahalanobis distance is found by using the current estimate of the range and range rate states to find a full state. This is then projected to the times of the two measurements and an expected measurement is generated. The expected measurement, $\hat{y}$, can then be compared with the input measurement, $\vec{y}$, and the error between these two states can be applied to the Mahalanobis distance equation shown in equation (4.22) where S is the measurement covariance matrix.

$$m = \sqrt{(\vec{y} - \hat{y})^T S^{-1} (\vec{y} - \hat{y})} \tag{4.22}$$

Four scenarios are investigated and the results are summarized in figures 4.15 through 4.16. Figure 4.15 shows the ratio of correct results, defined as all state components with less than one percent error. The Nelder-Mead algorithm more consistently converges to the true state in all

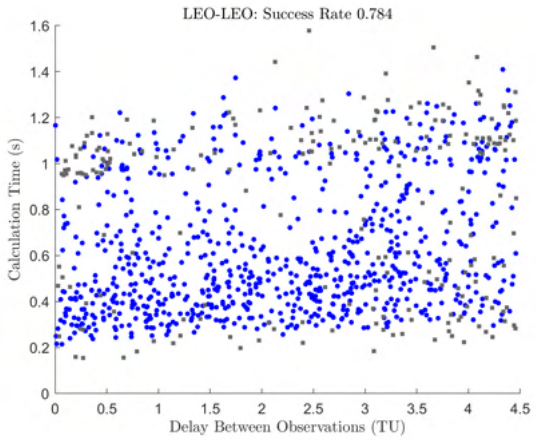

Figure 4.13: LEO observer to LEO target Monte Carlo particle swarm Results.

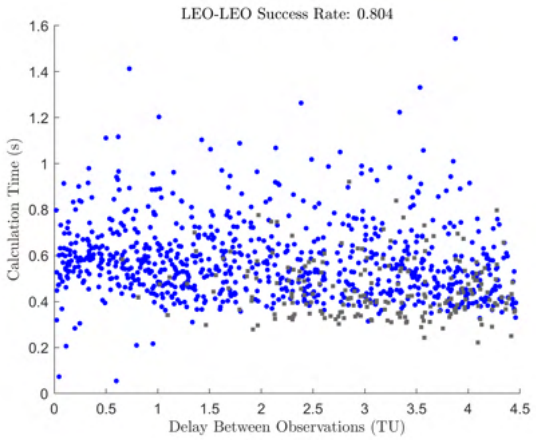

Figure 4.14: LEO observer to LEO target Monte Carlo Nelder-Mead Results.

cases. It's performance is worst with the LEO-LEO scenario. As shown in figure 4.14, this is due to the high error rate as the time between observations increases to the point that both objects have moved a significant fraction of an orbit. Figure 4.16 shows the calculation time distribution for all scenarios for the two algorithms. The box and whisker plots show that the Nelder-Mead algorithm is much more consistent in the time required to find a solution. The average computation time is lower for the Nelder-Mead algorithm.

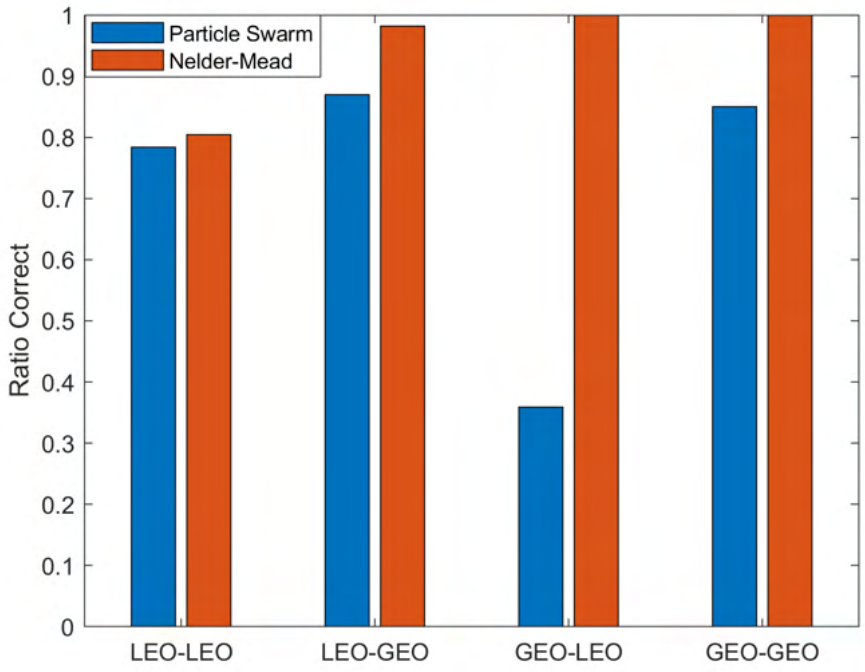

Figure 4.15: Ratio of correct solutions for both the particle swarm and Nelder-Mead approaches.

For the previous test cases with the Nelder-Mead algorithm, the maximum number of initializations per case is set to fifty. However, there is a trade-off between computation time and accuracy where the more times the algorithm is reinitialized with new randomly generated initial conditions, the more opportunities there are for the algorithm to converge to a global minimum and the more time there is potentially spent on each case. This trade-off is investigated by running the same set of one-thousand cases with several different values for the maximum number of runs. The results for both the accuracy of the solutions and the average time per case are shown in figures 4.17 and

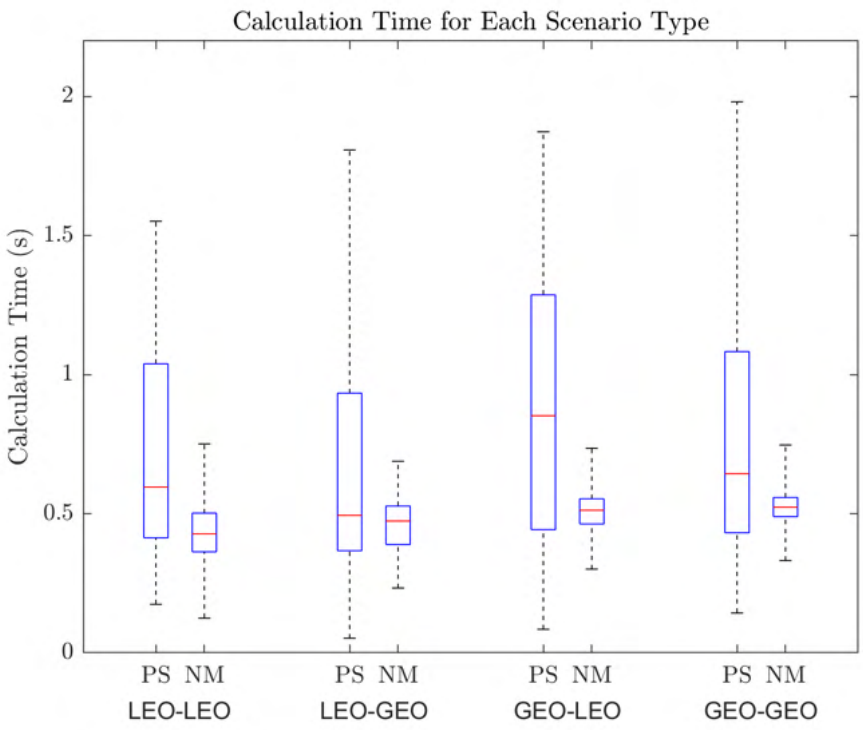

Figure 4.16: Particle swarm (PS) and Nelder-Mead (NM) computation time by scenario. The whiskers show the maximum and minimum while the box shows the twenty-fifth percentile, mean, and seventy-fifth percentile [1].

4.18 respectively. The accuracy shows a steep rise as the maximum number of runs increases up to fifteen followed by a sharp leveling off. After this point, more runs gradually increase the method's accuracy, but the change is under a percent between fifteen up to fifty. The time cost shows a much more linear result. The computation time increases steadily as the maximum number of iterations increases. The linear least squares fit for this line has a slope of 0.007844 seconds per run, and the coefficient of determination for the fit is 0.9902.

4.3.2 Multi-Orbit Scenario

All of the previous simulations for the developed initial orbit determination algorithms have assumed that the time between observations was less than a full period of either the observer or target object's orbit. However, if the time between observations is sufficiently large, it becomes feasible that the target or observer has completed more than a full revolution around the Earth.

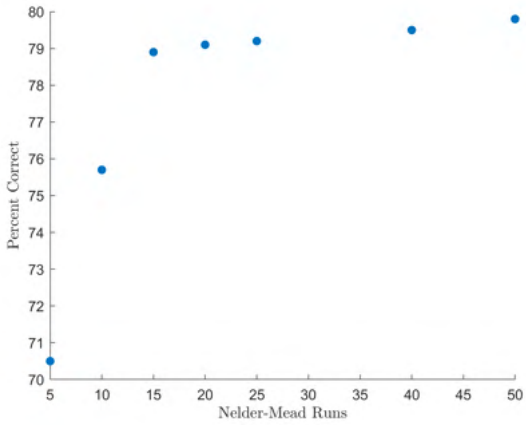

Figure 4.17: LEO observer to LEO target percent correct as a function of the number of Nelder-Mead runs.

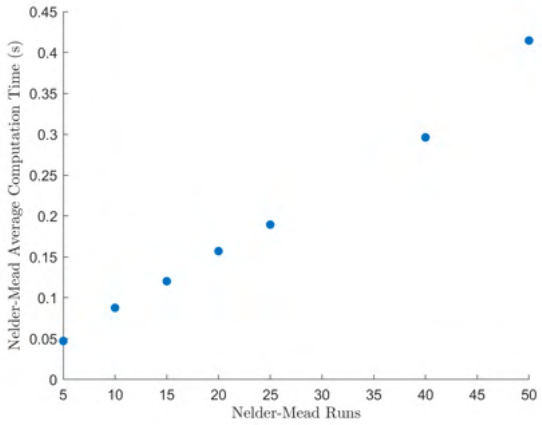

Figure 4.18: LEO observer to LEO target average computation time as a function of the number of Nelder-Mead runs.

The developed algorithms remain functional even in this case. There is no longer a unique solution for the case of two observations. Figure 4.19 shows the summed cost function contours for a LEO observer and LEO target case with a thirty-thousand second delay between the two angles and angle rates observations. The penalty function retains its shape, but the interior is divided into a series of curves. Each of these curves represents a feasible solution given the available data. The plot has several distinct non-trivial global minima, and only the correct solution is marked with a red dot. Accordingly, a single solution is not sufficient to characterize the set of global minima of this cost function.

By removing the ability to exit early when a sufficiently low Mahalanobis distance solution is found, the Nelder-Mead approach can be altered to report all identified minima, and the accuracy of the method can be measured by checking if the true state is in the set of reported solutions. With this modification, figure 4.20 shows the scatter plot results for a LEO observer and LEO target object with time delay between observations up to thirty-thousand seconds. The accuracy of the algorithm falls as the time delay increases.

4.3.3 Response to Measurement Error

One of the key metrics for determining the utility of an orbit determination algorithm is the accuracy of the estimated state as the measurement error increases. For an initial orbit determination method, the limited number of measurements fed into the algorithm prevent the error compensation and data smoothing that may be implemented in a precise orbit determination method such as a non-linear Kalman filter. The final state estimates will almost certainly deviate from truth with the inclusion of any measurement error. However, the distribution of these state estimates can be investigated to demonstrate the reliability of an IOD algorithm in response to the inclusion of measurement error.

In order to measure the reliability of the proposed IOD algorithm, a series of Monte-Carlo tests are performed for several states. These Monte-Carlo tests begin with a pair of exact measurements corresponding to a simulated orbit scenario. Gaussian-distributed noise is then added to these

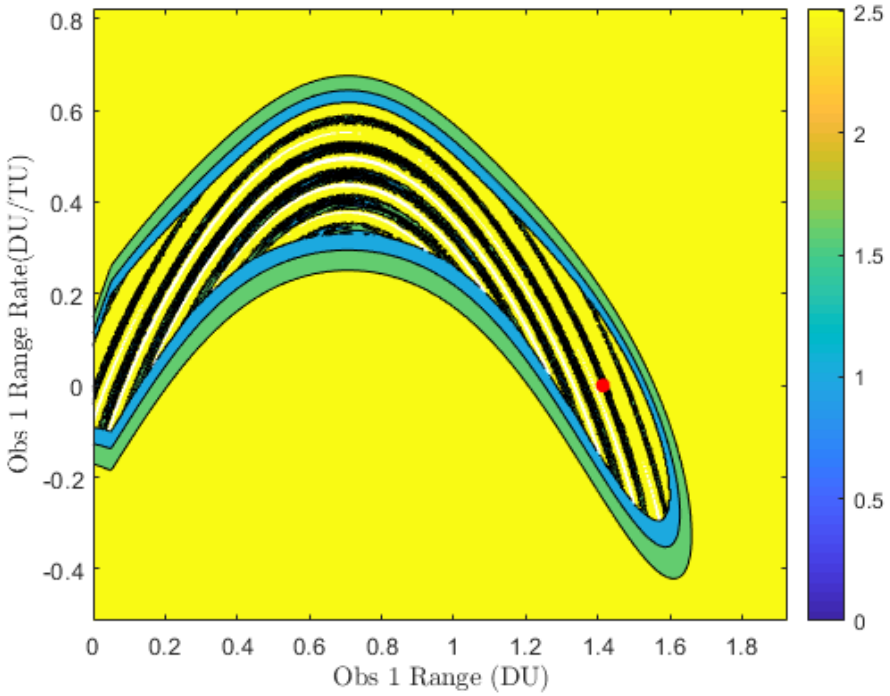

Figure 4.19: LEO observer to LEO target Multi-Revolution cost function contours.

measurements to model measurement error. The errors for each angle are treated as independent of each other, so the only cross covariance terms are between the angles and their respective angle rates. These erroneous measurements are then fed through the initial orbit determination algorithm and a prediction is found for the range and range rate between the observer and target. Repeating this process, the distribution of estimates represent the algorithm's response to measurement error with a set distribution. This process can be repeated for different observation scenarios and measurement error distributions in order to build up a holistic understanding of the algorithm's response to measurement error.

This process has been performed for the proposed IOD algorithm. The results for example scenarios are provided in figures 4.21 through 4.28. Two different measurement error standard deviations are explored. One case for angular standard deviations of 5 arcminute and angle rate standard deviations based on the error propagation of measuring the endpoints of a streak for a ten second integration. The other case for angular standard deviations of 1 arcminute and a

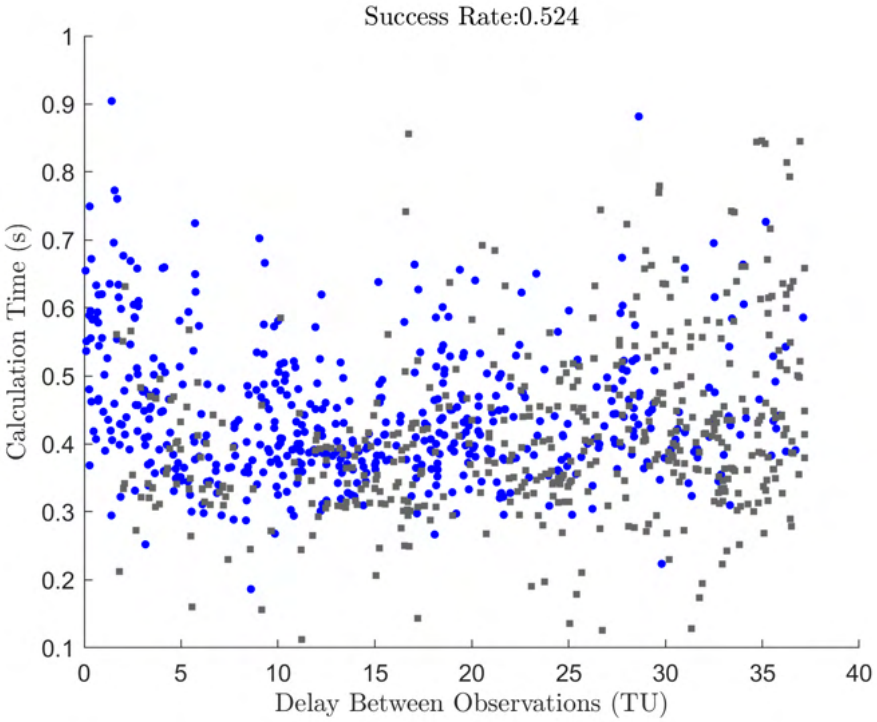

Figure 4.20: LEO observer to LEO target Multi-Revolution Monte Carlo method.

similarly derived angle rate standard deviation. Assuming no covariance between the two endpoint measurements, the propagated error for the angle rates is given in equation (4.23) where t is the image integration time. Using the streak start position as the angle measurement, the cross covariance terms between the angle measurement and the corresponding angle rate measurement is the negation of the angle variance divided by the integration time.

$$\sigma_{\dot{\phi}}^2 = \frac{\sigma_{\phi_1}^2 + \sigma_{\phi_2}^2}{t^2} \tag{4.23}$$

For the case of a LEO observing a LEO, the results for the tested case are shown in figures 4.21 and 4.22. The resulting distribution of state estimates is well-approximated by a Gaussian curve in the solution space. For the first case with a 5 arcminute angular standard deviation, the estimate distribution has a standard deviation of 480 meters along the range axis and 0.31 meters per second along the range rate axis. This drops to 20.5 meters and 18 millimeters per second when the measurement error is reduced to 1 arcminute.

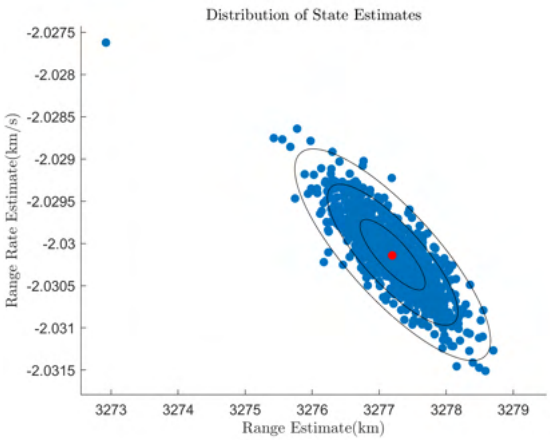

Figure 4.21: LEO observer to LEO target distribution of first position state estimates with 5 arcminute measurement error.

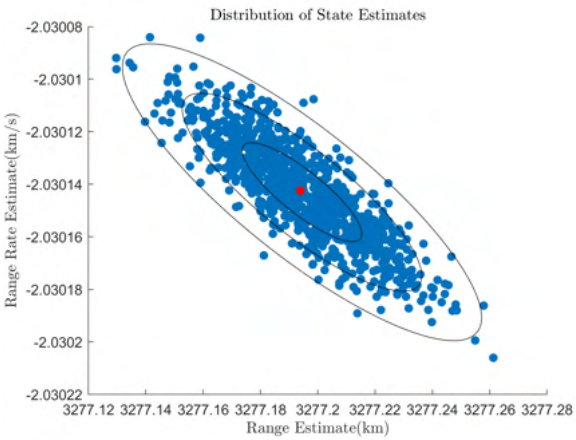

Figure 4.22: LEO observer to LEO target distribution of first position state estimates with 1 arcminute measurement error.

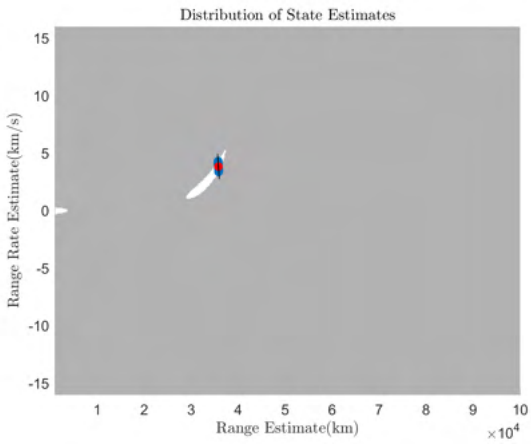

Figure 4.23: LEO observer to GEO target distribution of first position state estimates with 5 arcminute measurement error. Inadmissible regions are in gray.

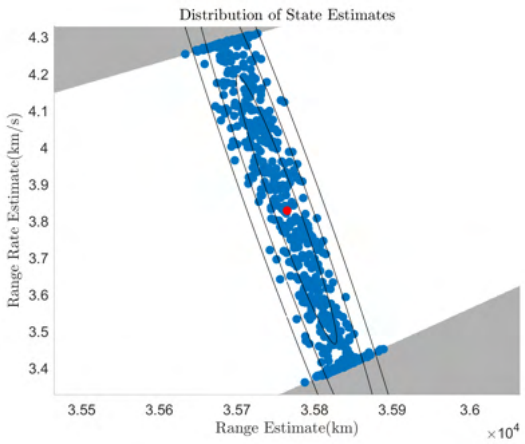

Figure 4.24: Figure 4.23 zoomed in to the region around the true state.

Moving to orbital regimes with at least one of the observer or target in a much more distant geosynchronous orbit, new features begin to appear in distribution of solutions. Figures 4.23 and 4.24 show the total distribution of one thousand estimated solutions to a case with a LEO observer and GEO target. The true solution is marked in red. While the general shape of the distribution is that of a two-dimensional Gaussian. Part of it is cut off by the admissible region boundaries. Points are instead pressed up against the border of the admissible region and the spread of the distribution is cut accordingly. This restriction provides the benefit more tightly containing the distribution of estimates around the truth rather than allowing them to spread out as they would otherwise without the admissible region restrictions. Restricting the angle measurement error to one arcminute for the same scenario as shown in figures 4.25 and 4.26, significantly reduces the size of the distribution. It shrinks enough such that the admissible region bounds no longer impact the shape of the distribution. The admissible regions also help contain the distribution to the region around its true solution. With the admissible region restriction turned off, local minima residing in the inadmissible space can attract estimates at higher measurement uncertainties. Scenarios with a geostationary observer and low-Earth target showed a similar response to measurement error as these cases.

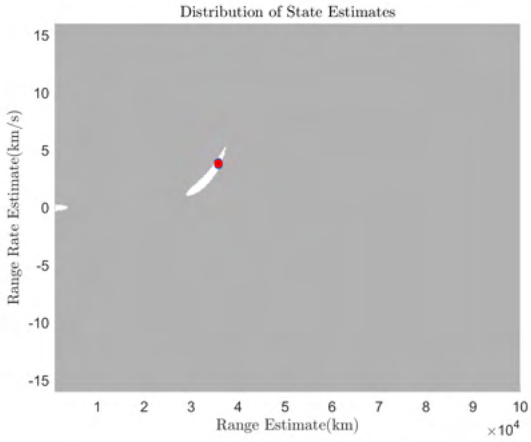

Figure 4.25: LEO observer to GEO target distribution of first position state estimates with 1 arcminute measurement error. Inadmissible regions are in gray.

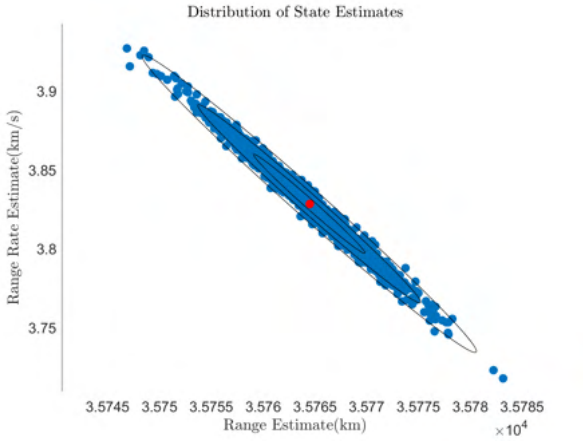

Figure 4.26: Figure 4.25 zoomed in to the region around the true state.

The state estimate error grows as the range between the observer and target increases and the relative velocity decreases. This is most evident in the results of a geosynchronous observer and target. While the majority of solutions still cluster around the truth, these cases report a larger number of false states with the same input error. Figures 4.27 and 4.28 show the distributions for the GEO-GEO case for angle standard deviations of 1 arcminute and 6 arcseconds respectively. These cases again show the fall off in state error as measurement error drops. A new feature is visible in this distribution as well. The region around the true solution has a long tail in the solution space. This is a result of the underlying contours having a long valley that curves through the state space. For GEO-GEO cases, this valley has a small derivative as one moves along the contour away from the solution. This makes other positions along the contour new likely minima as the measurements include more error. Notably, in all the presented cases, the true solution remains the most attractive point in the solution space, and the majority of all cases still fall into the region around the true solution. For the six arcsecond case, the estimates are the most dense around the true solution. This is most visible when focusing in on this region. Figure 4.29 shows this region and the relatively densely packed estimates along the minimal curve and around the truth. Most of the estimates that make up the tail visible in figure 4.28 are much less dense than this region. The six arcsecond case was used for the GEO to GEO case to show that the estimates do converge to the truth when the error is sufficiently small, as the one arcminute case has a wide distribution of range estimates not seen in the other scenarios. This is due to the large range between the observer and target that magnifies any angle measurement errors.

4.3.4 Comparison to Other IOD Methods

It has been established that the proposed IOD method differs from existing IOD methods in the types of inputs it accepts, its wide range of usable domains, and its approach to the IOD problem. What remains is a comparison between the proposed method and existing methods in terms of their accuracy when subjected to measurement error. To that end, several methods were compared using the methodology employed in the previous section to demonstrate the differences

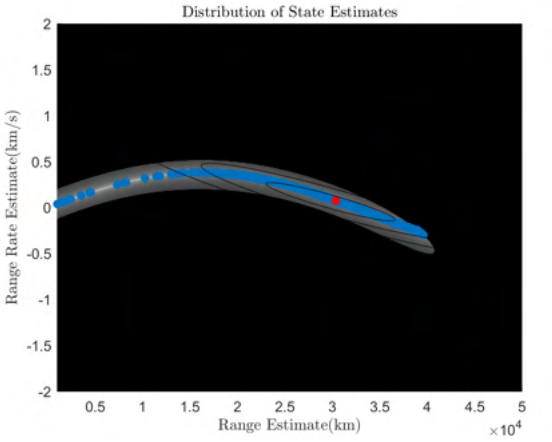

Figure 4.27: GEO observer to GEO target distribution and underlying contour of first position state estimates with 1 arcminute measurement error.

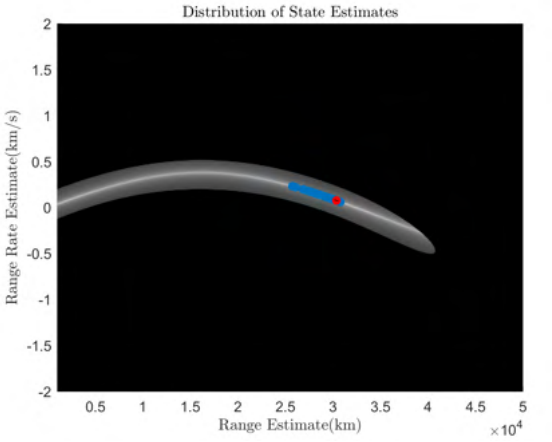

Figure 4.28: GEO observer to GEO target distribution and underlying contour of first position state estimates with 6 arcsecond measurement error.

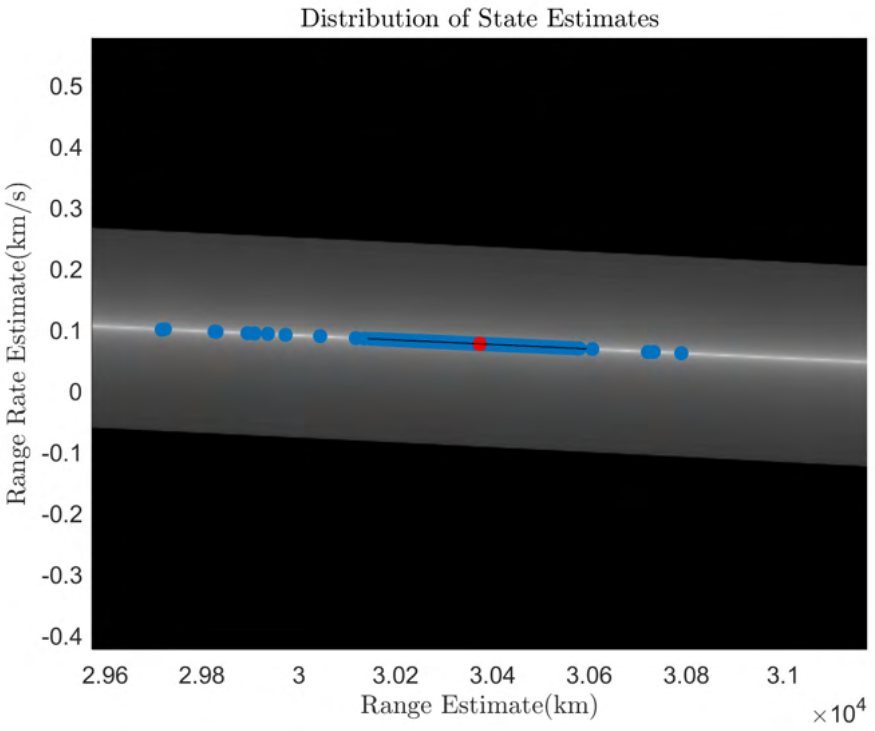

Figure 4.29: Magnified view of figure 4.28 around the true solution.

in results between the approaches.

Since the vast majority of existing methods are angles only, the comparison between the methods cannot perfectly match input data. Instead, the angles-only cases use two closely spaced measurements to simulate the endpoints of a streak and a third measurement from the start of the second streak. Since the practical implementation of the proposed method uses the endpoints of a streak to measure the angle rates, this implementation most closely approximates the input data to the proposed IOD algorithm.

4.3.4.1 L_n Method

The L_n method that is most similar to the developed method has limitations in its useful domain of observer and target orbits. The method does not work if the target's orbit is similar to or below the observer's orbit. Accordingly, in terms of the tested domains, the LEO-GEO case is the only scenario that provides a useful comparison to the proposed method. The two methods were

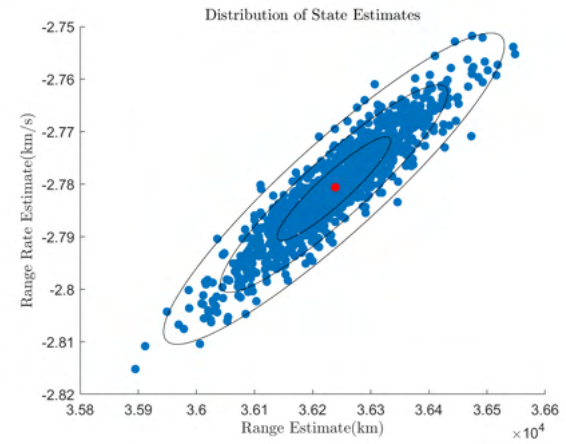

Figure 4.30: LEO observer to GEO target distribution for the L_n IOD algorithm.

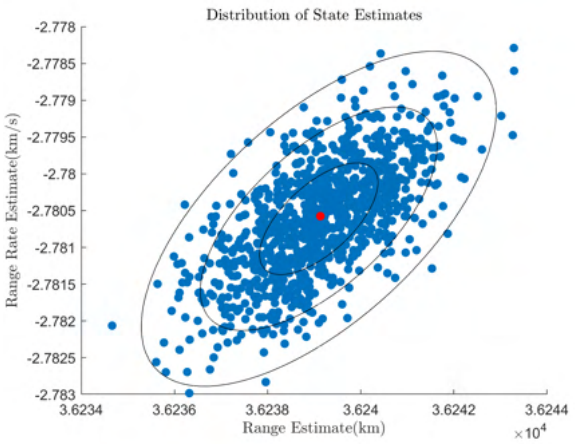

Figure 4.31: LEO observer to GEO target distribution for the proposed L_n+rates IOD algorithm.

compared with one arcminute measurement standard deviations, and the orbits were defined as two circular orbits. The observer was given an initial Earth-centered inertial position of [6000,2000,-200] kilometers and a rotation axis aligned with the vector [0.2,0.1,2]. The target has was given an initial Earth-centered inertial position of [42164,0,1000] kilometers and a rotation axis aligned with the vector [0.1,0.3,3]. The admissible regions applied in the proposed method restricted ranges to those below 6×10^4 kilometers and range rates with magnitude below 15 kilometers/s. In this case, the resulting distributions for one-thousand test cases are shown for L_n and the proposed method in figures 4.30 and 4.31 respectively. These proposed method has significantly tighter error margins for the state estimates in this case. For the L_n case, the range standard deviation is 102 kilometers and the range rate standard deviation is 11 meters per second. For the proposed algorithm, the range standard deviation is 39 kilometers and the range rate standard deviation is 0.8 m/s. It should be noted, that opening up the admissible region limits on the range allows several of the test cases to fall into local minima at much larger ranges than the true solution. These regions can feasibly be ruled out by observable magnitude constraints in many cases, and only a few of the one-thousand samples fell into these local minima.

4.3.4.2 Double-R

Escobal's Double-R IOD algorithm is more restricted then L_n+rates in terms of the domains in which it successfully finds solutions [80]. The algorithm requires a relatively accurate initial guess for the target's orbital radius at two of the measurement times as well. In order to compare Double-R with the proposed L_n+rates method, the two methods were tested for a LEO observing a satellite in medium-Earth orbit (MEO). For this test case, the observer was given an initial Earth-centered inertial position of [6000,2000,-200] kilometers and a rotation axis aligned with the vector [0.2,0.1,2], identical to the previous case. The target was given an initial Earth-centered inertial position of [12200,0,1000] kilometers and a rotation axis aligned with the vector [0.1,0.2,2]. Both of these were given velocities to form circular orbits. The initial guess for the Double R range of the first two measurements was provided as 12750 kilometers. The distribution of state estimates for both algorithms are provided in figures 4.32 and 4.33.

For the Double-R method, the range standard deviation is 2.45 kilometers and the range rate standard deviation is 4 meters per second. For the same physical scenario, the proposed method has a range standard deviation of 7130 meters and a range rate standard deviation of 1.27 meters per second. The proposed L_n+rates method has tighter bounds on its response to measurement error.

4.3.4.3 Gauss

Gauss's IOD method is tested against these algorithms as well under the same LEO to MEO scenario as the Double-R method. Similar to the other angle-only methods, Gauss's method does not work well unless the observer is inside the orbit of the target. Its use cases are accordingly limited for orbiting observers. The distribution of state estimates, shown in figure 4.35 is similar to that of the L_n method for the same scenario as shown in figure 4.34. Both perform worse than Double-R and the proposed L_n+rates method in terms of the breadth of the distribution of their state estimates. For Guass's method, the range standard deviation is 6.87 kilometers while the

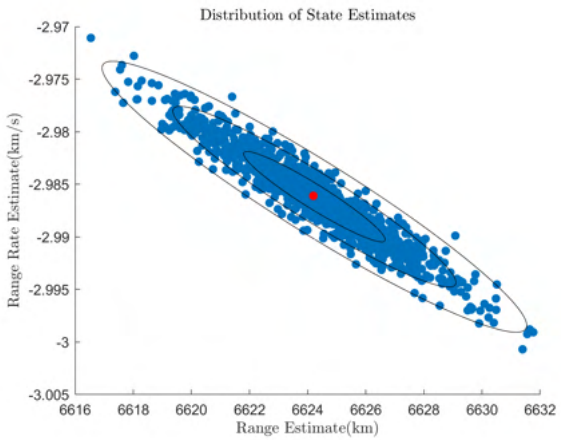

Figure 4.32: LEO observer to MEO target distribution for the Double-R IOD algorithm.

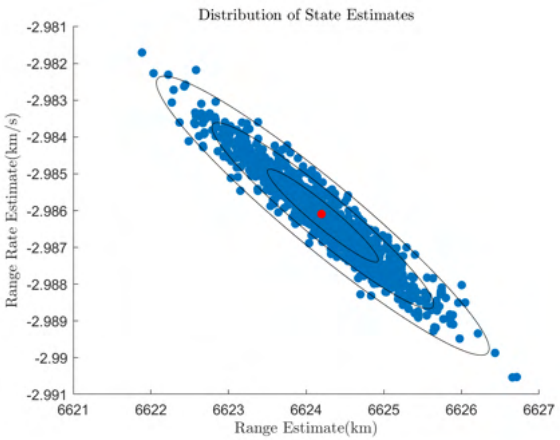

Figure 4.33: LEO observer to MEO target distribution for the proposed L_n+rates IOD algorithm.

range rate standard deviation is 13.2 meters per second. Compared to the results for the L_n+rates method, the distribution for Gauss is wider. The results for all tested IOD algorithms are provided in the table below and figure 4.36. The proposed approach is much more capable of consistently providing accurate state estimates than the other algorithms. However, this comes at the tradeoff of a much slower computation time as shown in figure 4.36.

4.3.4.4 Comparison Across Domains

The four methods were compared in terms of their ability to successfully find state estimates in the four combinations of GEO and LEO observers and targets. Each method was compared with the same set of randomly generated observer and target orbits. The Gauss method did not perform well, but this is partially due to the algorithm working best for short times of flight and with the observer inside the orbit of the target. The time of flight was randomly selected between 100 seconds and one hour. Many of the failures in the Gauss set are due to too long times of flight. Reducing the time of flight limits to 50 seconds to 500 seconds for the LEO-GEO case increases the Gauss success ratio to 43.6%.

For comparing the algorithms in all domains, the Double-R method was given the initial

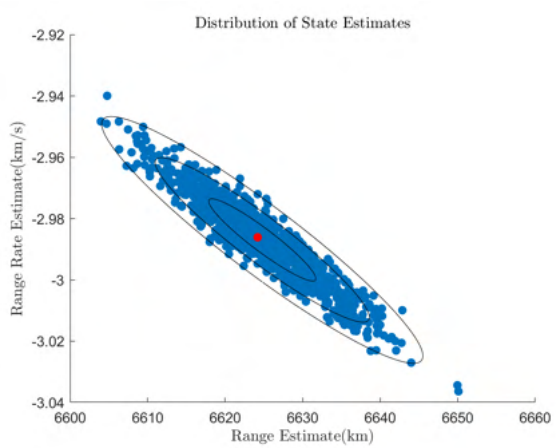

Figure 4.34: LEO observer to MEO target distribution for the L_n IOD algorithm.

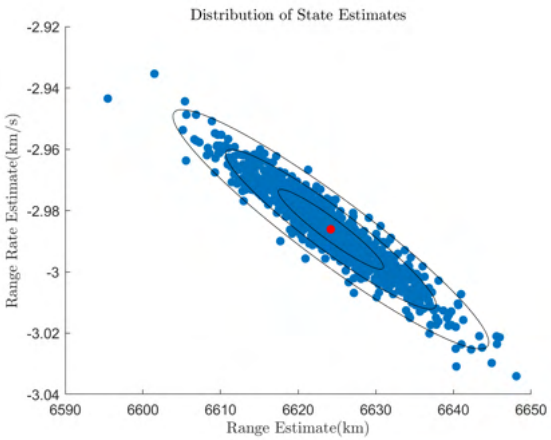

Figure 4.35: LEO observer to MEO target distribution for the proposed Gauss IOD algorithm.

Orbital Domain	Gauss	Double-R	L_n	L_n+rates
LEO-LEO	0%	77.6%	0%	99.6%
LEO-GEO	6.6%	0.3%	98.7%	100%
GEO-LEO	0%	37.7%	0%	100%
GEO-GEO	0.1%	8.1%	0%	99.7%
Ground-LEO	0%	70.8%	11.3%	99.9%
Ground-GEO	6.9%	25.2%	66.3%	100%

Table 4.2: Success ratio for each IOD algorithm and orbital domain for a 1000 sample Monte-Carlo test.

guesses recommended by David Vallado's *Fundamentals of Astrodynamics and Applications* of $r_1 = 2756.274$ km and $r_2 = 12820.05537$ km [18]. Accordingly, the algorithm fares much better in domains closer to the initial estimate. The algorithm performed best with LEO observer and target follow be GEO observer and LEO target. The two domains with GEO targets performed poorly.

L_n performs very well when the observer is well inside the orbit of the target as in the LEO-GEO case. However, the algorithm does not work well when the observer and target are in similar orbits or the observer is in a higher orbit than the target.

The computation time for all of the existing algorithms is much faster than the proposed L_n+rates algorithm. This is one of the major drawbacks of the proposed method if the computational resources are extremely limited or a large amount of data must be processed in a short time span.

4.3.5 Empirical Verification

To verify that the Nelder-Mead algorithm could accurately estimate the translational state of actual spacecraft. The algorithm is empirically verified on a series of images taken of actual Earth-orbiting space objects. The images are taken with the Omnidirectional Space Situational Awareness Telescope (OmniSSA) at the University of Colorado at Boulder. This imaging system is composed of four all-sky cameras with overlapping fields of view to improve the effective signal to noise ratio. The boresights of the cameras are fixed with respect to the ground, so non-geostationary space objects appear to streak through the frame for longer integration times. Two observation

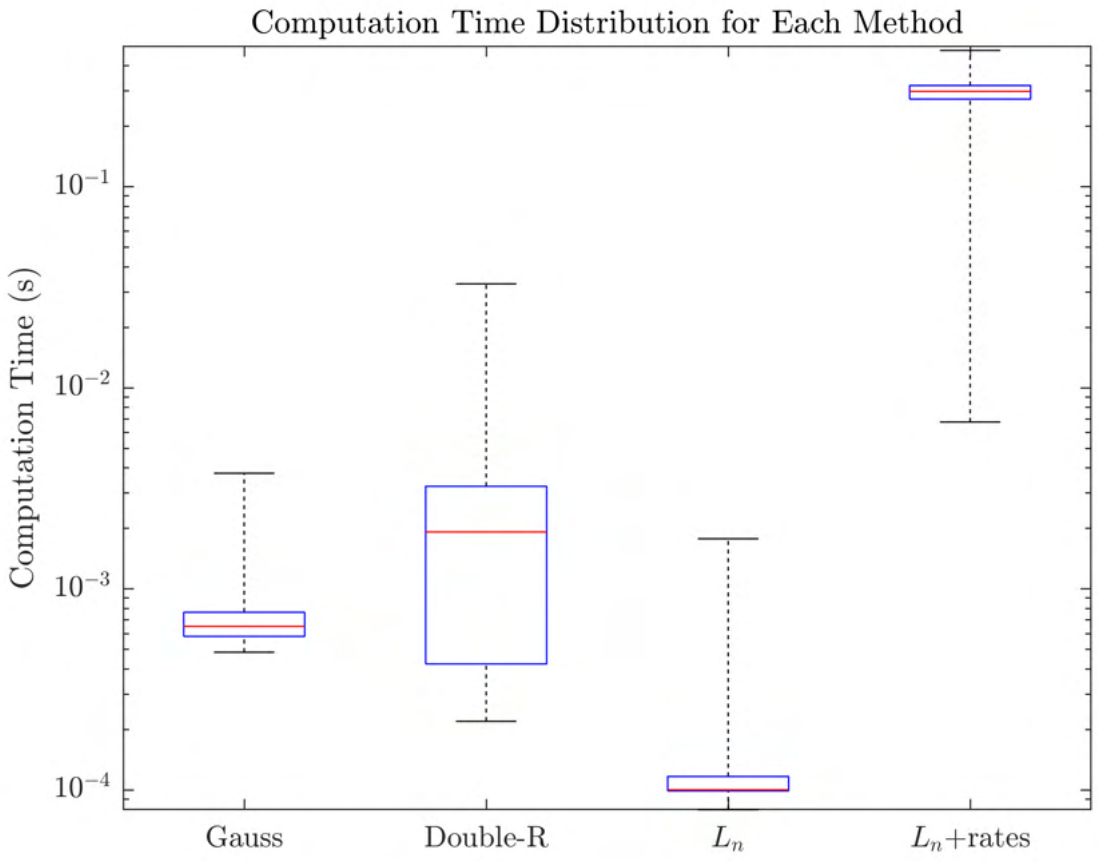

Figure 4.36: figure

Distribution of computation times for correctly estimated states for each IOD algorithm.

runs are included in the results. The first took place in the evening of October fourth, 2018, and the second was on November eighth, 2019. In the first of these sessions, the integration time was set to seven seconds. In the latter, the integration time was ten seconds.

To determine the true orbits of the imaged objects, the two-line element (TLE) catalogs for all unclassified objects are collected for the period around each observation date. This data set is then narrowed down by removing multiple listings for the same object taken at different times. The retained TLE is the listing with the closest epoch to the actual observation time. The catalog data is obtained from Space-Track [81]. In order to determine which TLE's the objects correspond to, the TLE data is passed into the open source night sky visualizer Stellarium, that uses an SGP4 integrator to show the motion of the input objects as the simulated time changes [82]. The motion of the simulated satellites is compared to the observed objects. Associations are made by comparing the simulated and observed predicted positions at the time of the observation as well as the type of object the TLE represents in order to ensure the TLE object is bright enough to reasonably create the observed streak.

The right ascension and declination of the target object at the start and end of each streak are then found by determining their respective precise locations in the image frame and passing that through a custom MATLAB function. This function uses the boresight angles determined by the Astrometry.net software image calibrator [55]. The lens distortion is accounted for with a LOESS-based calibration algorithm. Combined together, these tools allow a user to determine the right ascension and declination from the observed pixel coordinates in an imaged frame.

The right ascension and declination is translated into a heading vector through equation (4.24). The angle rates are determined by finding the heading for both the start and end of the streak and assuming straight line motion on the surface of the celestial sphere to determine the heading rate through numerical differentiation. The observer position is found by converting the latitude, longitude, and altitude of the observer to Earth-centered, Earth-fixed coordinates and pre-multiplying this vector by a rotation matrix that is based on the observation time to translate these coordinates to the Earth-centered inertial frame. The observer velocity in the inertial frame

can then be found via numerical differentiation through equation (4.25).

$$\vec{p} = [\cos\delta\cos\alpha, \cos\delta\sin\alpha, \sin\delta]^T \tag{4.24}$$

$${}^{\mathcal{N}}\dot{R} = \frac{[NR(t)][RN(t+\delta t)] - I_{3\times3}}{\delta t}[NR(t)]^{\mathcal{R}}R \tag{4.25}$$

A table of the observed and expected results for a series of identified objects is provided. The expected and estimated ECI state vectors are compared. The expected values are based on SGP4 propagation of the TLE's while the estimated values are derived via the proposed IOD method. The data was collected with an ultra-wide angle lens with a diagonal field of view of 115 degrees. This significantly limited the accuracy of the measurements. This is combined with very short arcs between the measurements that limit the state observability. However, this does demonstrate how a lower cost sensor setup can be used with the proposed method to perform initial orbit determination. The inherit error in the TLE, propagated error for the time difference between the TLE and observation, and the measurement error in OmniSSA all contribute to the error between the true state and the estimated state. While these errors vary on a case by case basis, the inaccuracies of the TLE are expected to contribute about at least a kilometer and potentially an order of magnitude more after propagating to the time of observation [83]. The error between the propagated and estimated states depend on the distance to the target. The geostationary target, Kondor-E 1, has a position error of 2493 kilometers and the four LEO's have position errors between ten and thirty kilometers. The high error in the Kondor-E 1 estimate is likely a reflection of both the TLE's age and the large distance between the target and observer that magnifies any measurement error. This along with a much shorter subtended arc for a GEO object that magnifies angle-rate errors.

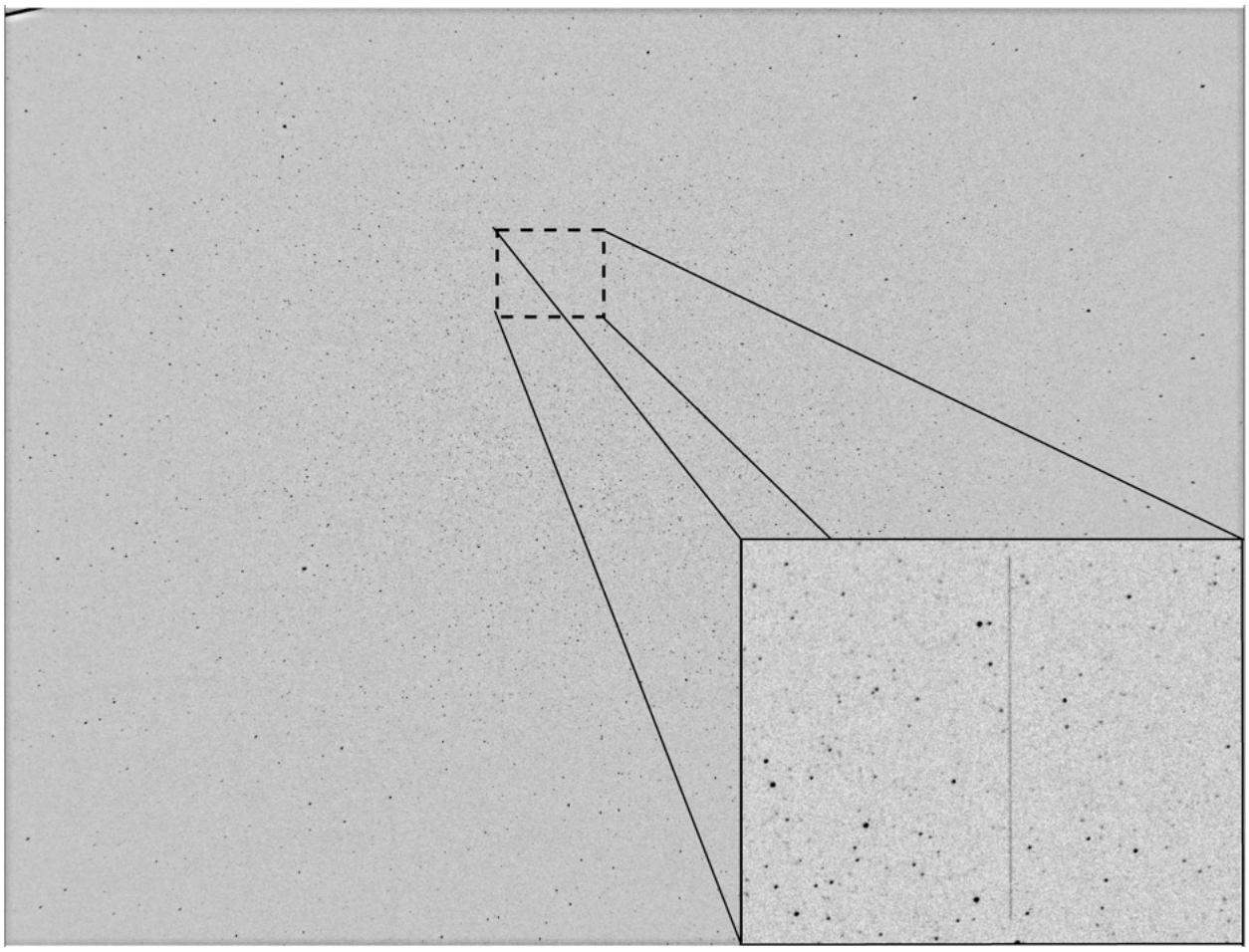

Figure 4.37: Streak generated by the Yaogan 3 satellite over a 10 second integration. Color-inverted and photometrically corrected image.

Target	TLE ECI Position	Estimated ECI Position	Pos. Error	TLE ECI Velocity	Estimated ECI Velocity	Vel. Error	M. Dist.	TLE Age	Orbit Fraction
Kondor-E 1	[30561,-29053,-96] km	[32473,-30632,-355] km	2493 km	[2.119,2.228,-0.00371] km/s	[1.917,.0902,0.00557] km/s	0.244 km/s	0.613	1.003 days	0.0136
Yaogan 3	[3977,-3741,4382] km	[3992,-3757,4397] km	26.967 km	[-4.308,2.217,5.787] km/s	[-4.357,2.039,6.224] km/s	0.475 km/s	7.944	0.210 days	0.00802
CZ-4C DEB	[4115,-3926,4136] km	[4132,-3944, 4140] km	25.308 km	[-4.045,2.092,5.994] km/s	[-3.671,1.525,6.399] km/s	0.791 km/s	1.868	0.741 days	0.00797
OAO 1	[2348,-5378,4092] km	[2350,-5397,4093] km	18.630 km	[6.753,3.170,0.284] km/s	[7.287,2.660,0.0926] km/s	0.763 km/s	7.935	0.053 days	0.0133
SL-3 R/B	[2839,-5187, 3557] km	[2845,-5196,3551] km	11.599 km	[-0.649,4.055,6.408] km/s	[-0.374,4.347,6.667] km/s	0.478 km/s	7.925	0.127 days	0.0140

Table 4.3: TLE and estimated values for observational verification with OmniSSA data.

4.4 Chapter Summary

A number of related developments on initial orbit determination have been formulated and discussed. First, a new method for initial orbit determination from angles and angle rates has been derived based on the L_n method developed by Reza Raymond Karimi and Daniele Mortari. This method was investigated by itself then abstracted into a cost function to allow traditional derivative-free optimization to find an estimate for the true state based on a pair of angle and angle rate measurements. Variations on this cost function that included admissible regions and the error in projecting the state into future times were investigated as well.

The capabilities of these methods as tools for initial orbit determination were then assessed under a variety of scenarios. These included different observer and target orbital regimes for every combination of Low-Earth orbiting satellites and near-geosynchronous satellites. This was investigated both by generating contours for the cost function of single scenarios as well as Monte-Carlo tests to quantify the algorithms' capabilities for randomly generated scenarios of each type. It was shown that the Nelder-Mead optimizer improves over the particle swarm in both computation time and its ability to converge to the correct solution. Furthermore, it was shown that by adding the two cost functions together, the resulting cost function developed a favorable geometry for efficiently converging to the correct solution for near coplanar scenarios. Finally, the Nelder-Mead optimizer over the novel cost function was shown to be capable of converging to the correct solution for multi-orbit scenarios. The proposed algorithm improves over existing methods in terms of orbit domain versatility and ability to reliably converge to close estimates of the truth. In over ninety-nine percent of tested single orbit scenarios, the proposed method yields an accurate solution. This comes at the cost of computation time where on the same hardware, the proposed method can take several deciseconds to yield a solution while other methods can return a result in several milliseconds or less.

Extending this work, the algorithm was modified and adapted to remove the assumption of two-body dynamics. This change was initially performed in order to adapt the algorithm for the

cislunar domain. However, the changes to the algorithm allow it to accept dynamics as an input and function as an IOD algorithm for N-body dynamics. These changes inspired a number of other advancements. Many of which could be pulled back into this original formulation to improve accuracy or better assess results.

Chapter 5

Initial Orbit Determination in Cislunar and Beyond

5.1 IOD Beyond Near Earth

In order to perform initial orbit determination beyond the reasonable limits of the two point-mass approximation, here a proposed method applies an optimization-based approach to find the state or set of states that can feasibly match the observations. Similar to the previous chapter, the approach leverages admissible regions to constrain the search space and improve convergence characteristics. An angles-only method is implemented, but the theory behind an angles and angle rates implementation is discussed as well. The optimization for the angles-only case is run over the four unobserved dimensions while the angles and angle rates case is run over the two unobserved dimensions. These are range and three velocity components for angles-only and range and range rate for angles with angle rates. The optimized cost function is the average of Mahalanobis distances between predicted and observed measurements based on a trial solution.

The approach implements JPL's SPICE Toolkit to accurately predict the relative positions of the Solar System bodies [84]. The gravitational influence of the Earth, Moon and Sun dominates the dynamics of orbits in cislunar space and beyond. The influence of the Sun is essential for identifying if a target object is in a cislunar, geocentric, or heliocentric orbit as targets sufficiently far from the observer can be difficult to differentiate in terms of orbit domains. The geopotential models of the Earth and Moon are also included up to fourth order using data from the EGM2008 and GRGM1200A respectively [85][86]. These include both the zonal and tesseral components.

The angles-only case uses three observations while the angles with angle rates case uses two

observations. The angles-only case is more applicable for larger range observations as the angular velocity of a target in cislunar orbit typically has a much slower inertial velocity than an object in a near-Earth orbit and the significant increase in separation distance between target and observer makes the observed angle rates much smaller. The relative drop in angle rate observability compared to the near-Earth orbits significantly restricts the types of measurements that can accurately capture angle rates. Accordingly, angles-only measurements are far more applicable for more distant domains. The cost function itself is the average of six Mahalanobis distances for angles-only observations and two Mahalanobis distances for angles with angle rates. Process noise is included in calculating the Mahalanobis distance so that the values accurately reflect the statistical likelihood of associating the measurements with a given state. From a single trial value, these are the distances between the predicted and observed measurements for the orbit propagated to each measurement time. The measurement treated as truth is then changed and the process repeated until each measurement is given equal weighting in the estimate. For perfect measurements and an accurate state guess, the cost function returns zero. We find that small measurement errors can still yield accurate estimated states.

The Mahalanobis-based cost function alone tends to form long curves through the solution space that some optimization methods may struggle with [45]. In order to improve the geometry of the underlying contours for optimization, modifications to the solution space and cost function are investigated. A method used to improve convergence for the cislunar domain is the modification of the admissible regions from the two-body implementation. The admissible regions for cislunar space differ substantially from the near-Earth orbits, and these limits are shifted accordingly to penalize the cost function in a way that improves convergence and limits the search domain. Investigated penalty function parameters include geocentric periapse and apoapse limits, heliocentric periapse and apoapse, obstruction of line of sight due to the Moon and Earth, occultation of the Sun on the target due to the Moon and Earth, and absolute limits on maximum and minimum velocities and ranges. Well targeted application of these admissible regions can significantly restrict the search space and improve the convergence properties of the optimization [87].

The geometry of the cost function is explored by generating contour plots over the solution space. The results of this initial orbit determination method are investigated in terms of convergence accuracy with respect to measurement error and the relative orbit geometry of the observer and target using a series of Monte Carlo simulations.

The proposed method is an approach for performing initial orbit determination beyond near-Earth space and the algorithm itself is neutral to the dynamics included in the propagator. The proposed method is a new approach that optimizes over a cost function to find feasible states that match observations. Implementations for the angles-only are investigated for relative accuracy along with the implementation theory for the angles and angle rates case. The necessary conditions to determine convergence is derived and implemented to exclude non-global minima from returning false solutions. The efficacy of these methods are demonstrated through a series of Monte Carlo simulations investigating the response of the algorithm across cislunar space and in response to measurement error.

5.2 Theory

5.2.1 Cost Function

At the core of the optimization-based approach is a cost function that models how well a trial solution fits the collected measurements. The cost function is a metric that should rigorously return a value that represents how well a state estimate fits the data considering the known measurement uncertainty. For this reason, the cost function is chosen as an average of Mahalanobis distances. The Mahalanobis distance is expressed in its most general form in equation 5.1 where $\vec{u}$ is the measurement vector, $\hat{u}$ is a predicted measurement vector based on the state estimate, and S is the measurement covariance matrix and is assumed known based on the particulars of the measurement system. For states propagated away from the reference state, the covariance matrix in the distance calculation includes uncertainty due to the propagation and process noise. The details of this implementation are developed in section 5.2.5. The Mahalanobis distance is computed in the

measurement space which depends on the observer's state. The observer's state is assumed known and any uncertainty folded into the measurement covariance matrix.

$$D = \sqrt{(\vec{u} - \hat{u})^{\intercal} S^{-1} (\vec{u} - \hat{u})} \tag{5.1}$$

Several measurements are required for the state to be observable. For angles-only measurements, three measurements are required for the state to be observable and for measurements with both angles and angle-rates, two measurements are required. Accordingly, the cost function is taken as the average of Mahalanobis distances relating each pair of measurements to each other both forward and backward in time. This process starts with a trial solution state at one of the measurement times. The choice of measurement time is arbitrary and it will be assumed that the first measurement time, t_1, is selected.

5.2.1.1 Angles and Angle Rates Measurements

With only two measurement times, the case with measurements containing both the relative angle from the observer to the target and the observed angle rates is covered first. In this case, the measurement vector contains four linearly-independent elements, two for the angles, and two for the angle rates. The solution space is two-dimensional. For this case, the cost function is the average of two Mahalanobis distances. The trial solution, a vector with range from observer to target and the range rate, is combined with the measurement vector for the t_1 measurement. This yields a full six-element state. This state is propagated to the time of the second measurement under the simulated cislunar dynamics to return an expected full state at the time of the second measurement. This can be combined with the known observer state to extract a predicted measurement at the time of the second measurement. The first Mahalanobis distance in the angles and angle-rates cost function is equation 5.1 with this propagated predicted measurement as $\hat{u}$ and the actual measurement for the second time as $\vec{u}$.

The second Mahalanobis distance is based on the first measurement. Since the state is built from the measurement at the initial time, directly calculating this distance would return zero and

there would be no compensation for the measurement error at t_1. To ensure this information is included in the cost function, the initial full state is propagated to the time of the second measurement. Then the unobserved parameters of the state at this time are extracted from the predicted state with the known observer state. This range and range-rate are then combined with the second measurement to build a full state estimate. This state can then be propagated back in time to the first measurement. At this point the approach becomes similar to that of the first Mahalanobis distance. The expected measurement is extracted and the Mahalanobis distance is built for the initial measurement using equation 4.22. These distances are added together to construct the cost function. This cost function is effectively the reciprocal of a likelihood function and the optimization process is equivalent to maximum likelihood estimation [88].

5.2.1.2 Angles Only Measurements

For more measurements, the required number of terms in the cost function sum is 2-permutations of n where n is the number of measurements. Accordingly, for the three angles-only measurements required to observe a full state, six Mahalanobis distances are included in the cost function. These are constructed similarly to those of the angles and angle-rates case. However, for angles-only data the measurement state has two linearly independent values and the solution space is four dimensional. The first two Mahalanobis distances take the estimated initial state and propagate it to the time of the second and third measurements. The expected measurements at these two times are compared to the observed measurements for the first two Mahalanobis distances.

The other four are generated in a similar manner as the second case in the angles and angle rates case. The initial state estimate is propagated to the second or third measurement time. Then the expected unobserved state is pulled from the propagated state and combined with the measurement at that time. This state is then propagated to one of the other measurement times to find an expected measurement to compare to the observation. Take $\vec{x}_i$ to represent the full six-element target state at measurement time t_i and $\vec{R}_i$ the six-element observer state at measurement time t_i. Let the function $\vec{x}_j = \vec{p}(\vec{x}_i, t_i, t_j)$, represent the propagation of the state at measurement

time t_i to measurement time t_j. Let the function $\hat{u} = \vec{f}(\vec{x}, \vec{R})$ extract the expected measurement from state $\vec{x}$ and let the function $\vec{v} = \vec{g}(\vec{x}, \vec{R})$ extract the unobserved states from state $\vec{x}$. Finally, let the function $\vec{x} = \vec{h}(\vec{v}, \vec{u})$ build a state $\vec{x}$ from a measurement, $\vec{u}$, and an unobserved state $\vec{v}$. With this, first two Mahalanobis distances are $D_{1,2}$ and $D_{1,3}$ and expressed by equation 5.2 while the latter four Mahalanobis distances are $D_{2,1}$, $D_{2,3}$, $D_{3,1}$, and $D_{3,2}$ where $D_{i,j}$ is given in equation 5.3.

$$D_{1,j} = \sqrt{\left(\vec{u}_j - \vec{f}\left(\vec{p}(\vec{x}_1, t_1, t_j), \vec{R}_j\right)\right)^{\mathsf{T}} S_j^{-1} \left(\vec{u}_j - \vec{f}\left(\vec{p}(\vec{x}_1, t_1, t_j), \vec{R}_j\right)\right)} \tag{5.2}$$

$$D_{i,j} = \sqrt{\left(\vec{u}_j - \vec{f}\left(\vec{p}(\vec{h}(\vec{g}(\vec{p}(\vec{x}_1, t_1, t_i), \vec{R}_i), \vec{u}_i), t_i, t_j), \vec{R}_j\right)\right)^{\mathsf{T}} S_j^{-1} \left(\vec{u}_j - \vec{f}\left(\vec{p}(\vec{h}(\vec{g}(\vec{p}(\vec{x}_1, t_1, t_i), \vec{R}_i), \vec{u}_i), t_i, t_j), \vec{R}_j\right)\right)} \tag{5.3}$$

5.2.2 Extension to Other Measurements

Angles-only and angles with angle rates measurements are possible types of observations with electro-optical systems. The approach can also be extended to radar measurements which return high accuracy values for range and range rate with lower accuracy estimates for the relative angles. Due to the higher uncertainty angle measurements, it would be reasonable to collect three observations or more to generate a state. However, only two measurements are required to generate a state. The method is similar to the angles-only case. There are 2-permutations of n projections where n is the number of measurements. The solution space spans over the angle rates or velocity components orthogonal to the displacement vector between the target and observer. Other than the methods needed to project the state space into the measurement or solution space, the fundamental approach to performing the algorithm with different types of measurements in unchanged.

5.2.2.1 Admissible Region

With only the cost function as described in the previous section, any optimization would be unconstrained and suffer from the associated penalties to computational efficiency as the optimizer attempts to search an unconstrained solution space. By applying known limits to the feasible solution states as a penalty function, the search space can be constrained and the algorithm can

receive the corresponding benefits of a smaller search space. The most basic limits that can be implemented are restrictions directly on the feasible values of the solution space. For measurements of angles and angle-rates these would be limits on the feasible range and range-rate between the observer and target. The range can be constrained by a combination of the known capabilities of the observing instrument and the maximum observable cross section of the target as a target sufficiently small and far away eventually becomes undetectable. The range-rate can be constrained by assuming physical limits on the velocity of the object depending on the types of objects intended to be characterized. For example, if the target is assumed to be bound to the Earth-Moon system, range-rate can be limited by the Earth escape velocity. If the measurements are angles-only, then the entire velocity component of the state is a part of the search space, and a similar limit can be applied.

The next type of constraint that can be applied as a penalty function are those that would have prevented detection in the first place. This includes line-of-sight constraints where detection would have been impossible due to the Earth or Moon occluding the target's position from the observer. The line-of-sight constraint can be implemented directly from the geometry of the system bodies and positions of the observer and estimated target position. Similarly, for visible-spectrum optical data, in most cases it can be assumed that the light coming from a target object that enables detection is light reflected from the Sun. Accordingly, a target object in the Earth or Moon's umbra would not be visible. Another geometric constraint can be applied for this case to remove solutions that would place the target object in the Earth or Moon's umbra.

Constraints regarding the orbit of the shape of the orbit itself can also be applied, but there is a cost-benefit to consider in their implementation for the cislunar domain. Under the two-body assumption, the orbits of a target object can be assumed to be conic sections and for non-escaping objects, the orbits can be assumed periodic. This is not necessarily the case for cislunar objects and other target objects acted upon by significant forces other than the Earth's gravity. Without the two-body assumption, constraining a target object's orbit based on future and past states requires a full propagation of the orbit to some user-selected time in the past and

future to detect violation of orbital constraints. Under the two-body assumption, cases such as Earth-impacting orbits, highly eccentric orbits, or escaping orbits can be detected directly from the the orbital elements that can be derived from a single full state estimate. These penalties can still be applied to objects under additional forces, but the implementation must balance the time-cost of an expensive propagation into the past and future to catch violation of these limits and the benefit of a more tightly constrained search space. An optimal solution will depend on assumptions regarding the target object and goals for the implementations speed of convergence.

Turning these various constraints into penalty functions first requires representing violation of these constraints as continuous values rather than Boolean. For search space limits, the violation can be simply implemented as the amount that the estimate violates the condition. For a line-of-sight penalty, the cost can be implemented as equation 5.4 where $\vec{r}$ is the estimated position state of the target, $\vec{r}_B$ is the position state of the occluding body, $\hat{p}$ is the pointing vector from the observer to the target, and R_B is the radius of the occluding body. For an umbra penalty, if the angle between the body-relative position state vector of the target and the target-relative position state vector of the Sun is less than ninety degrees, the penalty expressed in equation 5.5 can be applied. Otherwise, the target object is necessarily not in the umbra and the body. The $\vec{r}_\odot$ is the relative position vector of the Sun to the Earth.

$$f = \max\left[0, -\left(\vec{r} - \vec{r}_B\right) \cdot \hat{p} - \sqrt{\left(\vec{r} - \vec{r}_B\right)^2 - R_B^2}\right] \tag{5.4}$$

$$f = R_B^{-2} \max\left[0, R_B - ||\left(\vec{r} - \vec{r}_\odot\right) \times \left(\vec{r} - \vec{r}_M\right)||\right] \tag{5.5}$$

Each of these penalties may be then be weighted by a multiplicative factor to account for the different scales of the units involved. For example, an angle penalty of one radian is far more severe than a position penalty of one kilometer and the weights can account for these differences. After the penalties are balanced, they can be added together and the penalty can be applied as $\exp\left(\text{penalty}\right) - 1$ and added to the cost function.

5.2.3 Nelder-Mead Optimization

The optimization is performed with a Nelder-Mead algorithm. This is a gradient-free algorithm that moves the vertices of a simplex each iteration to find a minimum of the cost function [77]. For measurements of angles and angle-rates, the search space is two-dimensional so the simplex is a triangle. For angles-only measurements, the search space if four-dimensional, so the simplex is the 4-simplex. The Nelder-Mead optimizer is chosen because, it has been shown that for angles and angle-rates measurements applied to two-body dynamics, the Nelder-Mead algorithm outperforms a particle swarm in both speed and accuracy [45]. The Nelder-Mead algorithm also works well for low-dimension problems which applies for these two and four-dimensional searches [89].

There is no guarantee of uniqueness for the initial orbit determination solution under the additional dynamics of the cislunar domain. Accordingly, rather than searching for a single minimal solution, the goal of the optimization is to find the full set of feasible solutions for the given set of measurements. A set of solutions can then be further narrowed down by collecting additional measurements and eliminating incompatible states. In order to do this, the algorithm is run multiple times for a given set of measurements and the set of distinct feasible minima are reported. Two methods for initializing the Nelder-Mead algorithm are compared in terms of their effectiveness at providing a set of solutions that contain a close estimate to the true state. One is random sampling of the admissible region for generating initial simplex vertices. The other is dividing the admissible region into separate regions and initializing the simplex in each region once.

By constraining the initial simplex vertices to lie within the admissible region, it is known that the algorithm starts with a set of unpenalized hypotheses. This guarantees that the algorithm can converge to a state within the admissible region. The admissible region is defined by the physical limits of the system, so as long as the user-defined limits are valid, the solution will exist within the admissible region.

5.2.4 Orbit Instability

With the addition of the influence of the Moon's gravity, the dynamics in some orbital regimes become chaotic. Small differences in the real and measured initial state grow with time to that point that the propagated predicted state no longer well-approximates the truth. As a consequence, for these chaotic orbits, there exists a time horizon beyond which a an initially accurate state estimate cannot predict the actual physical state. This time horizon can be characterized by a Lyapunov exponent [90]. The Lyapunov exponent for a given orbit can be approximated with equation 5.6. Where t_0 and t_1 are an initial and final time, $\Phi(t_1, t_0)$ is the state transition matrix from t_0 to t_1, and the matrix norm is the spectral norm [91]. An in-depth discussion of this approach is available in Ren, Li, and Zheng's 2020 paper.

$$\lambda = \frac{1}{t_1 - t_0} \ln ||\Phi(t_1, t_0)|| \tag{5.6}$$

5.2.5 Accounting for Process Noise

Process noise is the result of unmodeled dynamics influencing the motion of the an observed object. Since they are unmodeled, their influence cannot be accounted for directly. Instead, the uncertainty of states increases as the state evolves. In a Kalman filter, this can be accounted for by adding uncertainty to the covariance at the end of the prediction step. This would be expressed by adding $\Gamma(t_{k+1}, t_k) Q_k \Gamma^T(t_{k+1}, t_k)$ to the covariance as shown in equation (5.7). There P is the state covariance matrix, $\Phi(t_{k+1}, t_k)$ is the state transition matrix from time t_k to t_{k+1}, Q is the process noise covariance matrix, and $\Gamma(t_{k+1}, t_k)$ is the integrated influence of the process noise on the state.

$$P_{k+1} = \Phi(t_{k+1}, t_k) P_k \Phi^T(t_{k+1}, t_k) + \Gamma(t_{k+1}, t_k) Q_k \Gamma^T(t_{k+1}, t_k) \tag{5.7}$$

For propagating a state forward in time and measuring its Mahalanobis distance between observed and expected, process noise can be included by expressing the covariance matrix as shown

in equation (5.8). R_k is the measurement covariance matrix at time t_k, H is the measurement matrix the projects the state into the measurement space, and C projects the measurement into the state space. For an angles-only measurement, C is expressed as shown in equation (5.9). There r is the trial range value, α is the right ascension measurement at t_k, and δ is the declination measurement at t_k. The full expression brings the uncertainty of the first measurement and the process noise forward in time to combine it with the measurement noise and build a more accurate model of the uncertainty of the next measurement.

$$S = H(\Phi(t_{k+1}, t_k)CR_kC^T\Phi^T(t_{k+1}, t_k) + \Gamma(t_{k+1}, t_k)Q_k\Gamma^T(t_{k+1}, t_k))H^T + R_{k+1}) \tag{5.8}$$

$$C = \begin{bmatrix} -r\sin\alpha\cos\delta & -r\cos\alpha\sin\delta \\ r\cos\alpha\cos\delta & -r\sin\alpha\sin\delta \\ 0 & r\cos\delta \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{5.9}$$

A different approach is required to account for process noise backwards in time. A smoothing algorithm is adapted for this purpose. The expression for the state covariance in the smoothing algorithm with process noise is given in equation 5.10 [73]. There the subscript index on the covariance matrices are the time index the covariance corresponds to, and the superscript index is the highest time index that is included in the smoothing.

$$P_k^l = P_k^k + S_k(P_{k+1}^l - P_{k+1}^k)S_k^T \tag{5.10}$$

$$S_k = P_k^k\Phi^T(t_{k+1}, t_k)\left(P_{k+1}^k\right)^{-1} \tag{5.11}$$

$$P_{k+1}^k = \Phi(t_{k+1}, t_k)P_k^k\Phi^T(t_{k+1}, t_k) + \Gamma(t_{k+1}, t_k)Q_k\Gamma^T(t_{k+1}, t_k) \tag{5.12}$$

To adapt this for the three measurement angles-only case with the second measurement as the reference, $l = k+1$. Only two measurements need to be included in any case for three measurements.

Backwards projections are either from the third to second measurement or from the second to first. In this case $P_{k+1}^{l} = P_{k+1}^{k+1}$ which is the just the covariance of the later measurement. For the IOD scenario, this is just based on the measurement error at that time so $P_{k+1}^{k+1} = CR_{k+1}C^T$. and $P_k^k = CR_kC^T$. With this, the backwards state covariance, P_k^{k+1} is as expressed in equation (5.13). The Mahalanobis distance is then expressed as shown in equation (5.16).

$$P_k^{k+1} = CR_kC^T + S_k(CR_{k+1}C^T - P_{k+1}^k)S_k^T \tag{5.13}$$

$$S_k = P_k^k\Phi^T(t_{k+1}, t_k)\left(P_{k+1}^k\right)^{-1} \tag{5.14}$$

$$P_{k+1}^k = \Phi(t_{k+1}, t_k)CR_kC^T\Phi^T(t_{k+1}, t_k) + \Gamma(t_{k+1}, t_k)Q_k\Gamma^T(t_{k+1}, t_k) \tag{5.15}$$

$$d = \sqrt{(\vec{u}_k - \hat{u}_k)^T(HP_k^{k+1}H^T + R_k)(\vec{u}_k - \hat{u}_k)} \tag{5.16}$$

It is worth noting that, depending on the types of process noise included, this may involving taking the Moore-Penrose inverse for singular matrices. This happens if the sum in equation (5.15) does not make an invertible matrix.

5.2.6 Determining Convergence

For a single initialization of the Nelder-Mead algorithm, it will either exit when the maximum number of iterations is exceeded or it converges to a minimum and sequential best vertices on the simplex change by less than some tolerance. For a solution to be feasible, it should be close to zero since a state that perfectly predicts the observed measurements will have a zero cost. Since the cost function is the average of Mahalanobis distances, it accounts for measurement error so long as the measurement covariance matrix accurately reflects the actual measurement error. It also includes the process noise in the measurements propagated away from the reference state that is being estimated. If the returned solution is large, it can be immediately rejected as a local minimum but not global minimum. The question then becomes, what is close enough to accept as a feasible solution. The simplest approach would be to consider the cost purely as the average of

Mahalanobis distances and set a threshold based on hypothesis testing and a target false-positive rate. However, this does not account for the impact of the propagation of uncertainty. To include that, the propagator can be modified to find the state transition matrix by computing the Jacobian of the dynamics at each step in the integration. The resulting state transition matrix can be applied to a vector of the error from the propagated state at the propagation start time to find the error from the dynamics at the propagation end time. This can be used to analytically express the cost function as a function of the initial state estimate and the measurements.

Let $\hat{x}^*$ be the true dynamical state for the set of measurements, and let $\vec{k}$ be the vector of measurements with some covariance matrix, $P_{k,k}$. The calculated cost function can then be written as J, a scalar function of $\hat{x}^*$ and the mean of the measurement vector $\vec{u}_k$. Then one can find the Hessian of the cost function with respect to the six measurements for angles-only data or eight measurements for angles and angle-rates. With the Hessian of the cost function, it becomes feasible to take a second order Taylor expansion of the cost function. That second order expansion is expressed in equation 5.17. Start with the Taylor expansion of the cost function J around the solution. Since the cost function is assumed to be at the solution, J and $\frac{\partial J}{\partial \vec{x}}\big|_{\hat{x}^*}\delta\hat{x}^*$ should be zero. Then just looking at the expansion with respect to the measurement states simplifies to equation 5.18. Assuming that the $\vec{k}$ is normally distributed about $\vec{0}$, the mean of J with respect to $\vec{k}$ will be the expectation value of the second term [92].

The variance of J can be found similarly. Starting with the expectation value representation in equation 5.21, the result can be simplified to 5.23. The simplification is derived in Holzinger, Scheeres, and Alfriend's 2012 paper.

$$\begin{aligned} J(\hat{x}^* + \delta\hat{x}^*, \vec{\mu}_k + \delta\vec{k}) =& J(\hat{x}^*, \vec{\mu}_k) + \frac{\partial J}{\partial \vec{x}}\bigg|_{\hat{x}^*}\delta\hat{x}^* + \frac{\partial J}{\partial \vec{k}}\bigg|_{\vec{\mu}_k}\delta\vec{k} + \\ &+ \frac{1}{2}\delta\hat{x}^{*\intercal}\frac{\partial^2 J}{\partial \vec{x}^2}\bigg|_{\hat{x}^*}\delta\hat{x}^* + \frac{1}{2}\delta\vec{k}^{*\intercal}\frac{\partial^2 J}{\partial \vec{k}^2}\bigg|_{\vec{\mu}_k}\delta\vec{k} + \delta\hat{x}^*\frac{\partial^2 J}{\partial \vec{x}\partial \vec{k}}\bigg|_{\hat{x}^*\vec{\mu}_k}\delta\vec{k} + ... \end{aligned} \tag{5.17}$$

$$J(\hat{x}^*, \vec{\mu}_k + \delta\vec{k}) = \frac{\partial J}{\partial \vec{k}}\bigg|_{\vec{\mu}_k}\delta\vec{k} + \frac{1}{2}\delta\vec{k}^{*\intercal}\frac{\partial^2 J}{\partial \vec{k}^2}\bigg|_{\vec{\mu}_k}\delta\vec{k} = \nabla J\delta\vec{k} + \frac{1}{2}\delta\vec{k}^{*\intercal}J_{k,k}\delta\vec{k} \tag{5.18}$$

$$\mu_J = \mathbb{E}\left[\frac{1}{2}\delta\vec{k}^{*\intercal}J_{k,k}\delta\vec{k}\right] \tag{5.19}$$

$$\mu_J = \frac{1}{2}\operatorname{Tr}\left(J_{k,k}P_{k,k}\right) \tag{5.20}$$

$$P_{j,j} = \mathbb{E}\left[\left(J(\delta\vec{k}) - \mu_j\right)^2\right] \tag{5.21}$$

$$P_{j,j} = \mu_j^2 - \mu_j \operatorname{Tr}\left(J_{k,k}P_{k,k}\right) + \frac{1}{4}\operatorname{Tr}\left(J_{k,k}P_{k,k}\right)^2 + \frac{1}{2}\operatorname{Tr}\left(J_{k,k}P_{k,k}J_{k,k}P_{k,k}\right) \tag{5.22}$$

$$P_{j,j} = \frac{1}{2}\operatorname{Tr}\left(J_{k,k}P_{k,k}J_{k,k}P_{k,k}\right) \tag{5.23}$$

With the mean and variance of the cost about the solution found, the necessary condition for convergence is that the proposed solution has a cost that belongs to this distribution. A low-cost, but imprecise test for this is to decide in advance on a maximum number of standard deviations from expected to accept and only accept a result if the cost falls within that limit as defined by the derived mean and standard deviation. This would only require a single cost estimate, and the algorithm could exit immediately when an acceptable solution is found. A more precise test would involve running the optimizer multiple times and collecting those returned solutions that cluster about a value. A statistical hypothesis test could then determine if these solutions fit the expected distribution.

5.3 Results

5.3.1 Implementation

In order to assess the capabilities and limitations of the proposed algorithm, the approach is verified against a set of simulated scenarios. These scenarios model both an observer and target with a high fidelity orbital model that includes the gravitational influence of the non-dwarf planets, the Sun, and the Moon. As well as the geopotential model of the Earth up to order 360 based on the EGM2008 model [85] and the geopotential model of the Moon up to order 165 based on the GRGM1200A model [86]. Next, the simulated orbits included the influence of the solid tides of both the Earth and Moon. There are also non-gravitational forces included, these are the atmospheric drag and solar radiation pressure on the satellites.

The dynamics of the system used in the IOD algorithm itself were simplified beyond this model with various forces toggle-able to measure the costs and benefits of simplifying the model in various ways. The results of this investigation are discussed in section 5.3.3. The forces included in the IOD algorithm propagator are the point-mass influence of the non-dwarf planets, Sun and Moon and the geopotential model of the Earth and Moon up to order four.

The test cases were randomly generated along specific orbital regimes. Each observer regime and target regime pair was generated with two hundred cases. The observer regimes are equatorial geosynchronous orbit (GEO), equatorial medium Earth orbit (MEO), equatorial low Earth orbit (LEO), and L1 halo orbits both in and out of plane. The target regimes are a lunar near rectilinear halo orbit (NRHO) based on the planned lunar gateway, L1 halo orbits both in and out of plane, L2 halo orbits both in and out of plane, Molniya orbits, and highly eccentric Earth orbits (HEO). Not every pair of observer and target is considered. Rather interesting cases are chosen from the set of generated test cases.

For each scenario, simulated observations are generated for different time steps between observations and with different levels of measurement noise added in. The measurement noise is Gaussian and symmetric on the unit sphere. Trial cases can be selected by choosing a regime scenario, time interval between measurements, and measurement error.

5.3.2 Cost Function Structure

First, the structure of the solution space passed through the proposed cost function is explored. Previous work has shown that a similarly derived cost function applied solely to the near-Earth domain under two-body dynamics can yield a smooth and continuous contour within the admissible region when the measurements include both angles and angle rates [45]. In cislunar space, the more complex dynamics of the additional forces create additional opportunities for issues to arise with the search space. This could potentially result in multiple global minima for a given scenario or an unfavorable geometry for optimization. In order to determine if these issues occur, contours are generated in the search space for several orbit scenarios. These contours show the

behavior of the cost function as a function of the trial state.

Cases for the angles-only scenarios are generated. This has a four-dimensionsonal solution space, and contours can only clearly show the response to changes in two dimensions. As a result, angles-only contours are generated as a series of slices through the solution space in order to show the behavior of the cost function along all four search dimensions.

5.3.2.1 L4

The contours for the a target object in a stable orbit around the L4 Lagrange point have a relatively simple structure. The contours for angles-only measurements with the correct velocities orthogonal to the line-of-sight direction are shown in figure (5.1a). The structure of the contour shows a long valley of low cost for increasing range. This appears due to the ambiguity in the range state as three angles-only measurements taken in relatively quick succession for a distant object have a limited ability to discern the range. Near to the observer, several local minima appear. These represent nearby orbits that could approximate the observed measurements, though they do not perfectly mach as the costs do not fall to zero.

Deviations to the line-of-sight orthogonal velocities result in the contours shown in figures (5.1b) through (5.1c). As the orthogonal velocities move away from truth, the cost minima do not fall as deep and the main valley containing the true solution shifts. This indicates that the contour should be optimizable, though there remains the potential for an optimizer to return local minima if initialized to close to another minimum.

5.3.2.2 Low Lunar Orbit

Figure (5.2a) shows the range and range rate dimensions of the angles-only contour for a target object in low Lunar orbit as seen by an observer in GEO. The other two states are set to the true values. The red dot indicates the location of the true solution. The contours shown are the logarithm of the cost.

Figures (5.2b) through (5.2c) show the contour angles-only contours with the non-plotted

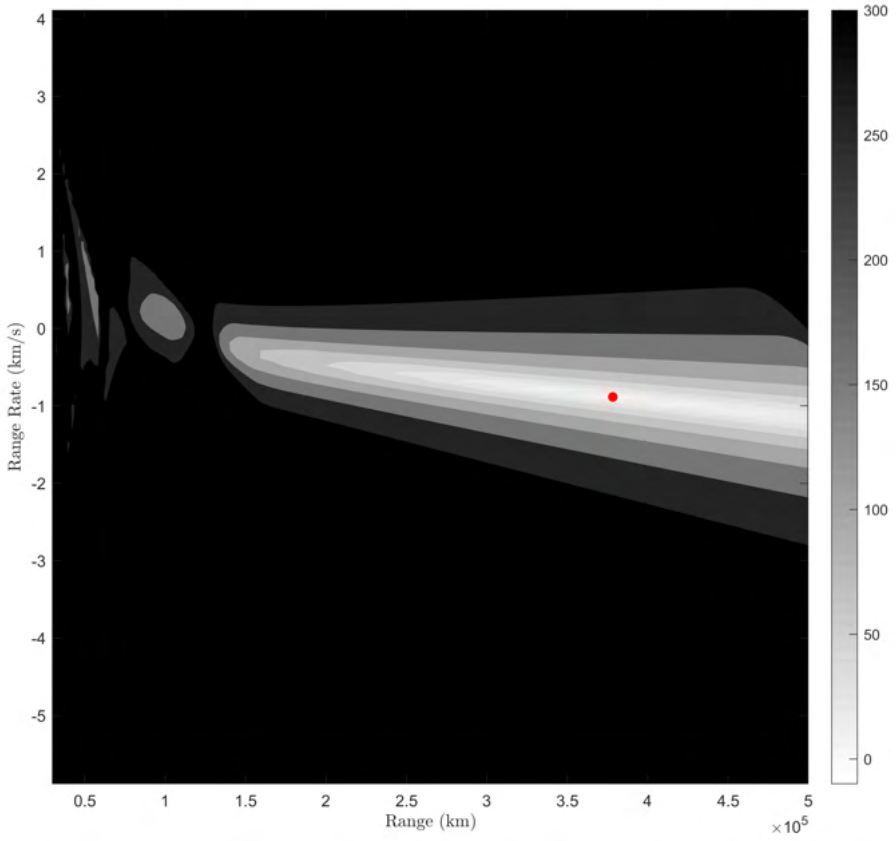

Figure 5.1a: Angles-only GEO observer to L4 target log(cost) contours with correct line-of-sight perpendicular velocities.

dimensions shifted slightly. This exposes the behavior of the solution space along the velocity axes perpendicular to the line of sight. The region of high cost near the solution reflects Moon-impacting orbits that trigger additional penalties on the cost function. There are also additional local minima in the nearby region such as the region at a higher range on the opposite side of the Moon-impacting lobe. This indicates that there is a significant potential for an optimizer to return minima that are not the solution. These minima also have a relatively low cost, so a relatively low amount of measurement error could make the minima indistinguishable in terms of detecting which is the global minimum and true solution. Comparing figure (5.2a) to figures (5.2b) through (5.2c) shows how these contours shift in the non-plotted dimensions. The nearby states are similar with slight shifts to the shapes of the contours and an increased cost around the solution minimum.

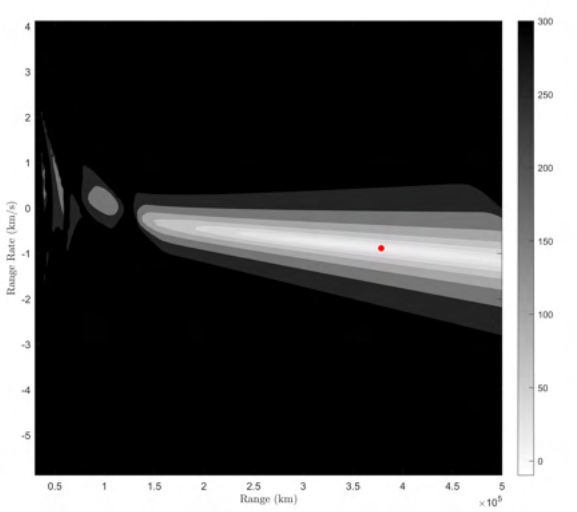

Figure 5.1b: Angles-only GEO observer to L4 target log(cost) contours with out of plane velocity increased by 0.03 km/s.

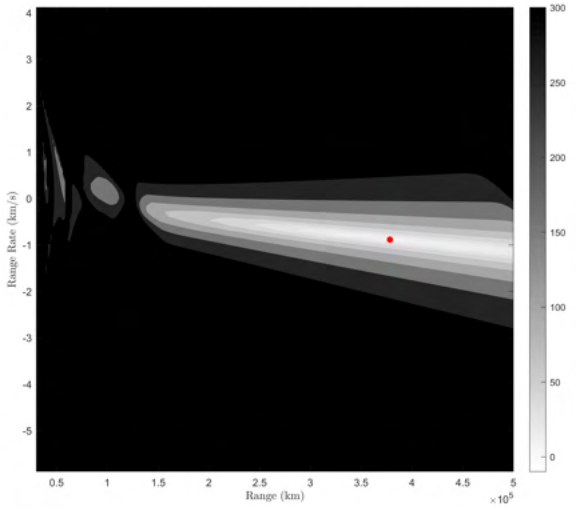

Figure 5.1c: Angles-only GEO observer to L4 target log(cost) contours with in plane velocity increased by 0.03 km/s.

This indicates that an optimization algorithm should be expected to fall into the solution if in the correct region rather than a set of solutions curving through the solution space.

5.3.3 Impact of Propagator Fidelity

The orbital propagator used to compare measurements to predictions at different times significantly impacts both the algorithm's computational efficiency and its accuracy. The optimal balance between these parameters will vary based on the goals of the implementation, but understanding the trade-offs with the propagator capabilities is valuable for appropriately choosing the implementation. The propagator is tested by initializing it with the first state in each orbit scenario and measuring the position error and required calculation time after one solar day of motion. Each regime is tested independently to isolate the influence of different propagator elements on different orbital regimes. The final position errors for the full fourth order geopotential model with the inclusion of all major Solar System bodies is shown in figure (5.3a). The distribution of calculation times is shown in figure (5.3b). Most orbits are modeled with less than a single kilometer error after a day of propagation. The halo orbits in particular perform well with the Lagrange point

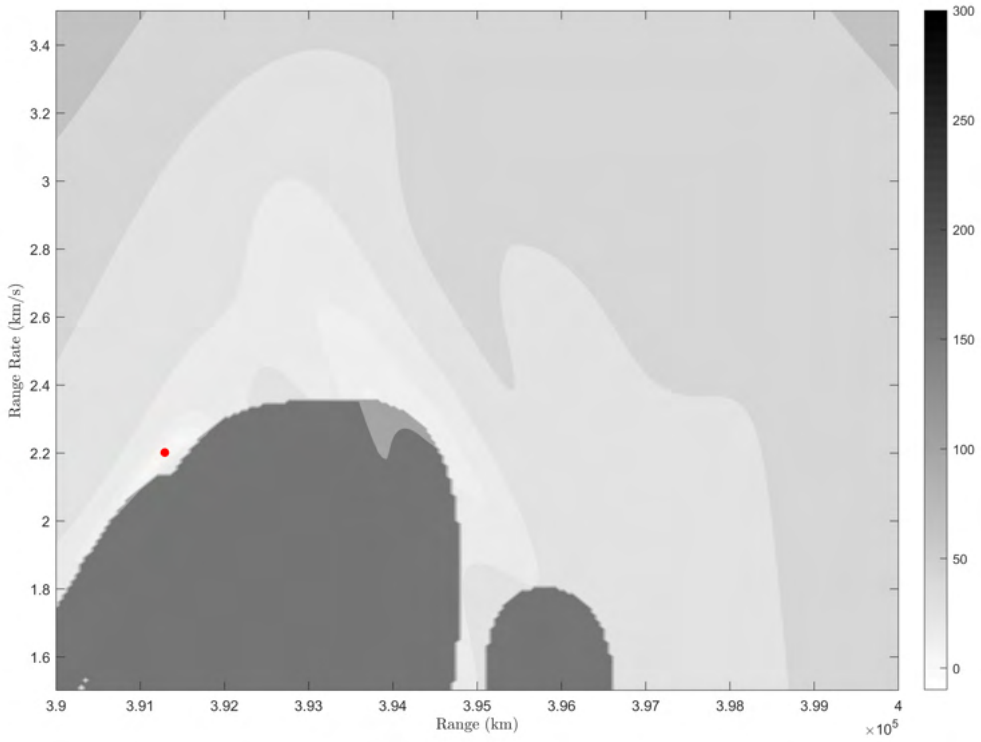

Figure 5.2a: Angles-only GEO observer to low lunar orbit target log(cost) contours with correct line-of-sight perpendicular velocities.

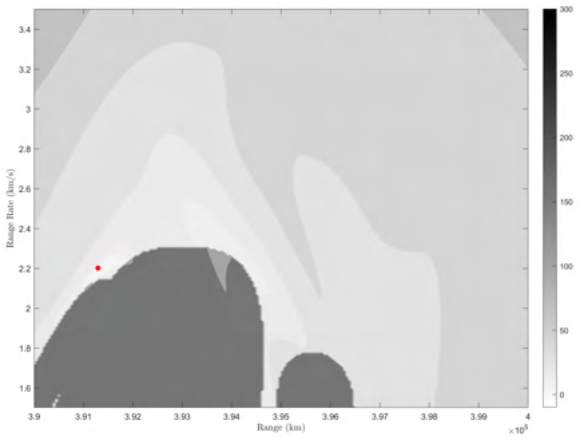

Figure 5.2b: Angles-only GEO observer to low lunar orbit target log(cost) contours with out of plane velocity increased by 0.03 km/s.

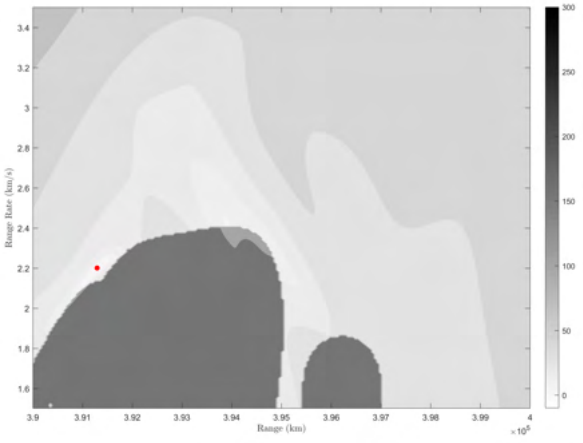

Figure 5.2c: Angles-only GEO observer to low lunar orbit target log(cost) contours with in plane velocity increased by 0.03 km/s.

halo orbit position error ending up at around twenty-five meters. The highly elliptical and Molniya orbits have the cases with the largest error. This is due to the unmodeled, higher orders of the

geopotential model, solid tides, and atmospheric drag.

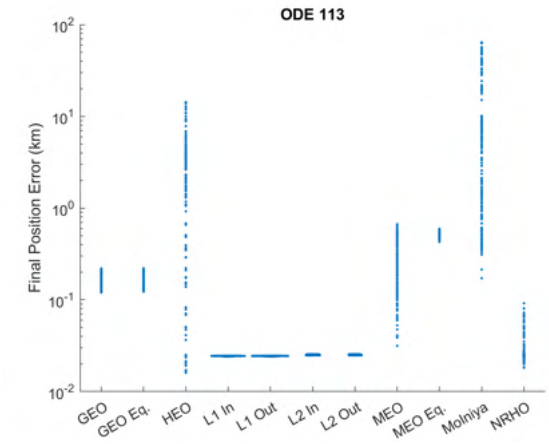

Figure 5.3a: Final position error logarithmic swarm chart for a series of orbital regimes after orbit propagation over a solar day using ODE 113.

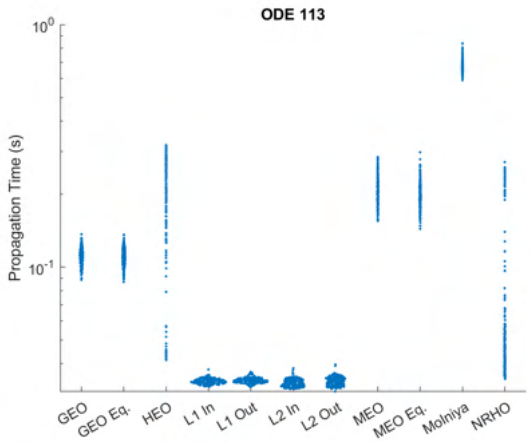

Figure 5.3b: Computation time logarithmic swarm chart for a series of orbital regimes for orbit propagation over a solar day using ODE 113

The previous case used MATLAB's ODE113, a nonstiff differential equation solver. This is compared against ODE45. The position error and time cost for this differential equation solver is provided in figures (5.3c) and (5.3d) respectively. While the accuracy of the two methods are similar, the ODE113 case performed significantly better in terms of the time cost of the propagation. This performance improvement is visible in all tested cases.

Figures (5.3e) and (5.3f) show the results after removing the influence of all bodies save for the Earth, Moon, and Sun. The time cost of the algorithm improved slightly, but there was a significant penalty to the final position state accuracy. Near-Earth MEO orbits were largely unaffected, but near rectilinear, L1, and L2 halo orbits saw over an order of magnitude loss of final state position accuracy to just under single kilometer level error for the Lagrange points and slightly more for the NRHO.

Reducing the order of both the Earth's and Moon's geopotential models to two yields the results shown in figures (5.3g) and (5.3h). The change increased the position error in the orbits closest to Earth by a small amount and had little observable impact on the more distant orbits. The impact was less severe than the loss in accuracy when the influence of other planets was removed. In

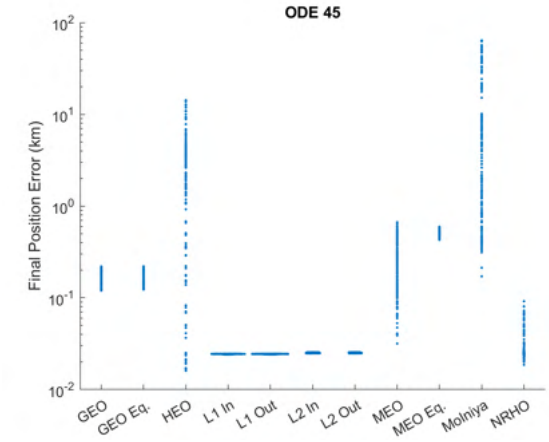

Figure 5.3c: Final position error logarithmic swarm chart for a series of orbital regimes after orbit propagation over a solar day using ODE 45

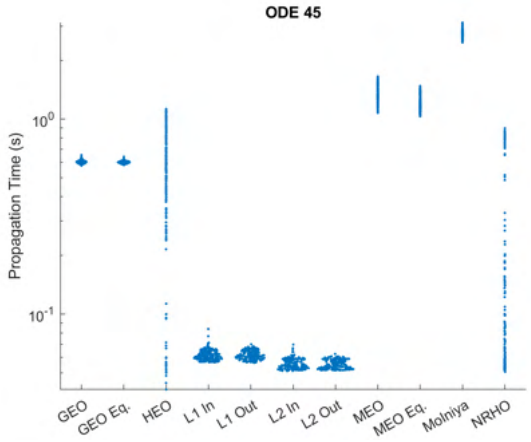

Figure 5.3d: Computation time logarithmic swarm chart for a series of orbital regimes for orbit propagation over a solar day using ODE 45.

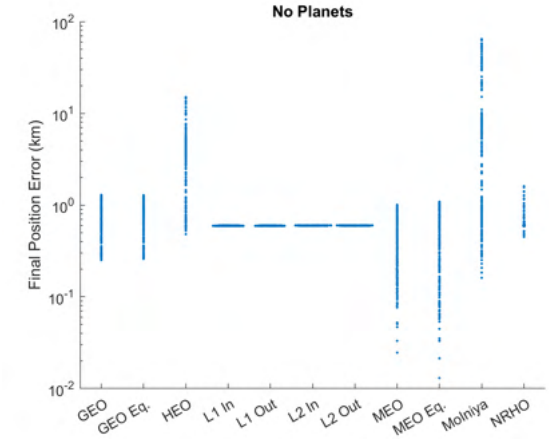

Figure 5.3e: Final position error logarithmic swarm chart for a series of orbital regimes after orbit propagation over a solar day. Propagated without the influence of planets outside the Earth system.

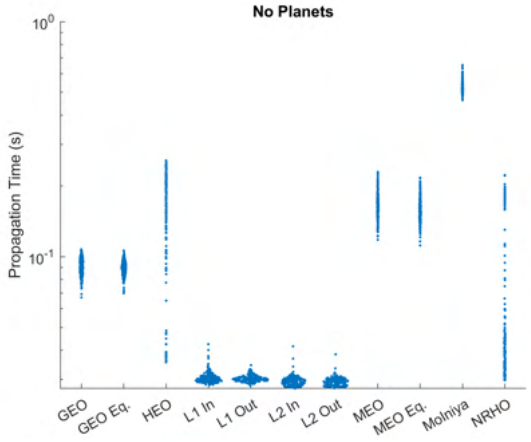

Figure 5.3f: Computation time logarithmic swarm chart for a series of orbital regimes for orbit propagation over a solar day. Propagated without the influence of planets outside the Earth system.

terms of the computation time, the time cost fell slightly across the board, though not by as much as removing the influence of the planets. Cutting the third and fourth order components of the geopotential model could be a reasonable method for saving computation time in the propagation step without significantly impacting the algorithm's accuracy.

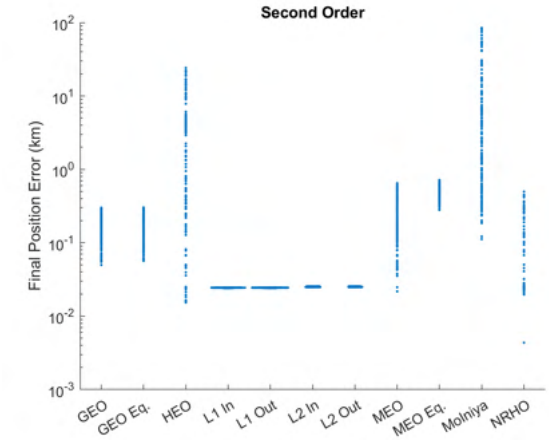

Figure 5.3g: Final position error logarithmic swarm chart for a series of orbital regimes after orbit propagation over a solar day. Propagated with second order geopotential models of the Earth and Moon.

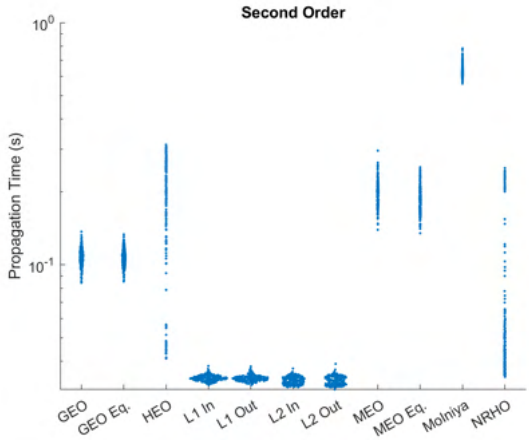

Figure 5.3h: Computation time logarithmic swarm chart for a series of orbital regimes for orbit propagation over a solar day. Propagated with second order geopotential models of the Earth and Moon.

Figures (5.3i) and (5.3j) show the propagation results when only the first order gravitational influence of the Sun, Earth, and Moon is included. While the computation time falls significantly for all cases, so does the position accuracy. All modeled GEO's and most modeled MEOS have final position errors over ten kilometers with some MEO's approaching just under 100 kilometers of error. The Lagrange point halo orbits are the only orbits with error consistently below one kilometer, but still perform far worse than the twenty-five meter error shown in the fourth order model.

5.3.4 Inclusion of Process Noise

The inclusion of process noise in the Mahalanobis distance shifts the calculated values and can be compared to approaches that do not include the process noise by the impact on the resulting metric. Since noise is added into the calculation, it would be expected that the Mahalanobis distance would decrease with the inclusion of process noise. How the distance changes in response to measurement noise is investigated for several different scenarios.

The process noise is set up to model solar radiation pressure as that is included in the

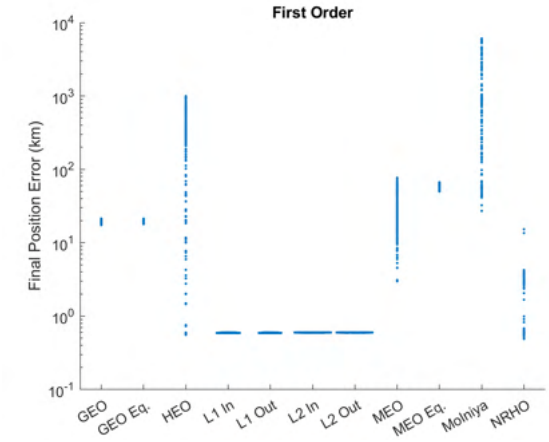

Figure 5.3i: Final position error logarithmic swarm chart for a series of orbital regimes after orbit propagation over a solar day. Propagated with only the point source influence of the Sun, Moon, and Earth.

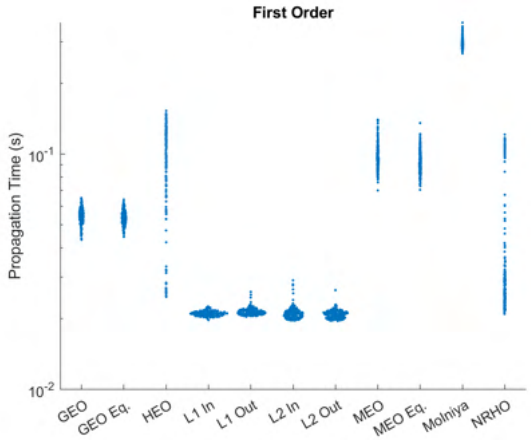

Figure 5.3j: Computation time logarithmic swarm chart for a series of orbital regimes for orbit propagation over a solar day. Propagated with only the point source influence of the Sun, Moon, and Earth.

simulated truth but not modeled in the optimizer's propagator. This is only one type of unmodeled dynamcis in the process noise. There are other forces such as higher order geopotential model terms that are not modeled in either the dynamics or process noise. Q is the scalar variance of the solar radiation pressure. Γ is defined as equation (5.24) [73]. For short time of flight, Φ and B can be treated as constant and the equation simplifies to equation (5.25). B is expressed in equation (5.26). x_s, y_s, and z_s are the relative ECI coordinates of the Sun from the target. r_s is the distance between the target and Sun.

$$\Gamma = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) B(\tau) d\tau \tag{5.24}$$

$$\Gamma = \Phi_k B(t_k)(t_{k+1} - t_k) \tag{5.25}$$

$$B = \begin{bmatrix} 0 & 0 & 0 & \frac{-x_s}{r_s} & \frac{-y_s}{r_s} & \frac{-z_s}{r_s} \end{bmatrix}^T \tag{5.26}$$

Fifty cases were randomly generated for several orbital scenarios and times. These are Geostationary observer to L1 halo in-plane target, L1 halo out-of-plane target, and NRHO. As well as low Earth orbit to Geostationary. For each of these scenarios, the original cost without process noise and the adjusted cost with process noise were found. The ratio of original cost over adjusted

cost was recorded and plotted in figure 5.4. Each scenario was investigated at four different times between observations. The results are plotted in a swarm charts for both the ratios and costs.

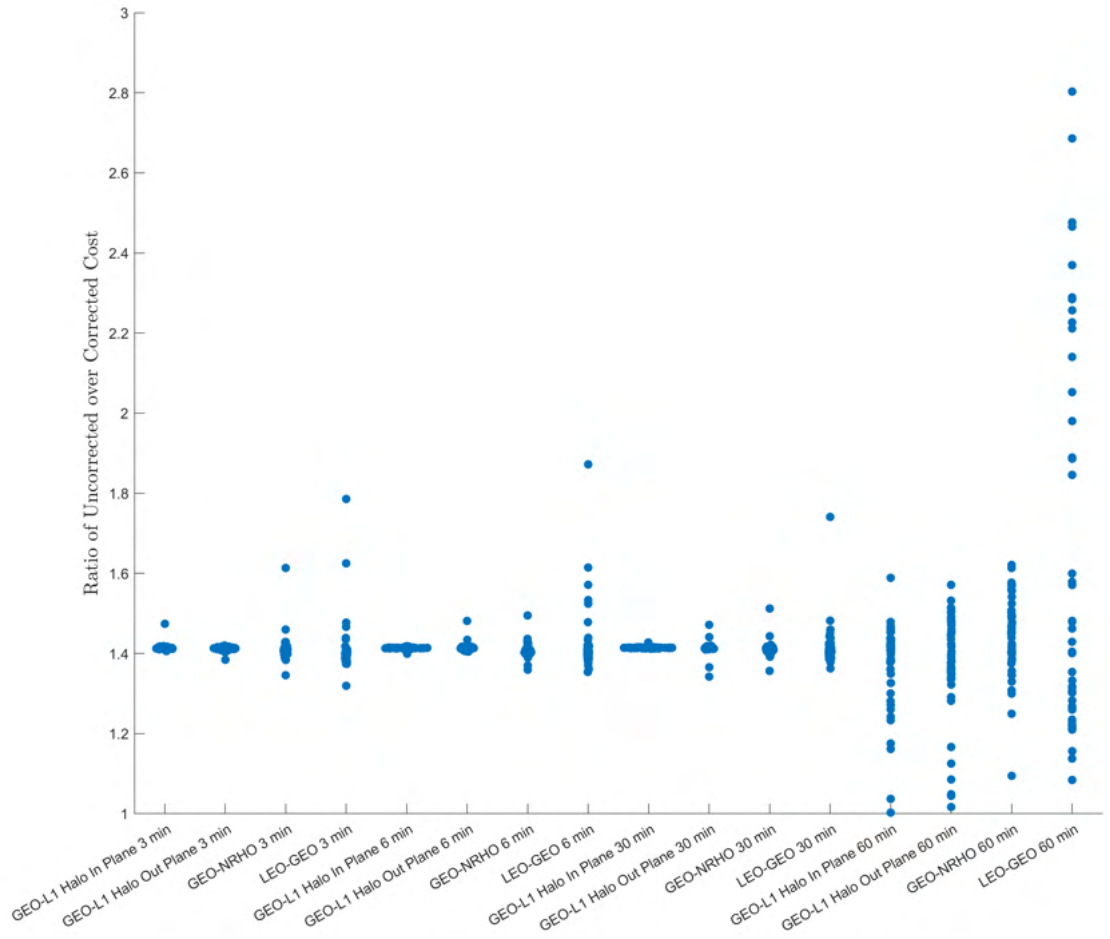

Figure 5.4: Ratio of cost with process noise over cost without process noise for four orbit scenarios at several times between observations.

The average cost ratio for all cases tended to hover around 1.4. Many of the results hovered around this value for short times of flight in the L1 halo target cases. When testing different amounts of process noise to apply, this value changes. The consistent multiplicative impact of including process noise is likely a result of the the way that the process noise is added. The additional term added to the measurement covariance can be fairly consistent for short times of flight. Larger times of flight spread out further with a much larger distribution when the time of

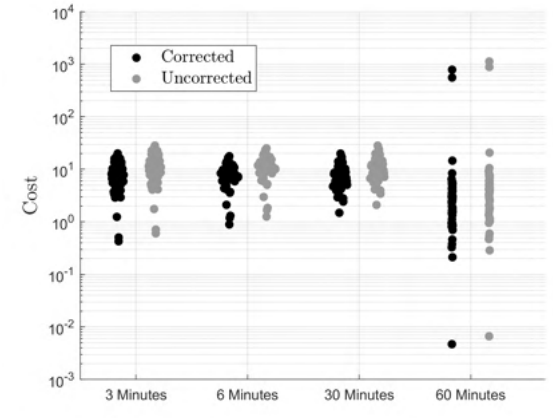

Figure 5.5a: Geostationary to L1 halo in plane cost before and after process noise correction.

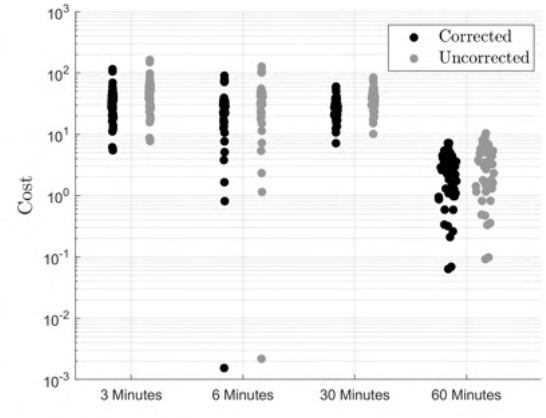

Figure 5.5b: Geostationary to L1 halo out of plane cost before and after process noise correction.

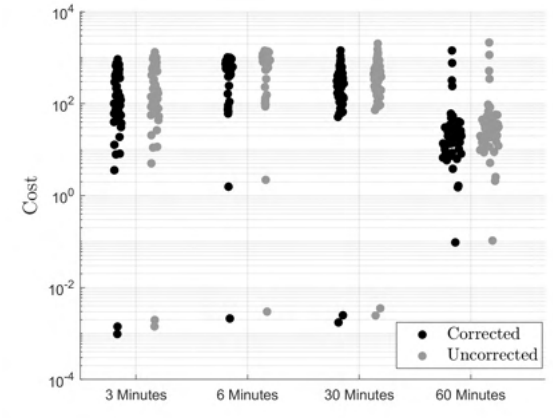

Figure 5.5c: Geostationary to NRHO cost before and after process noise correction.

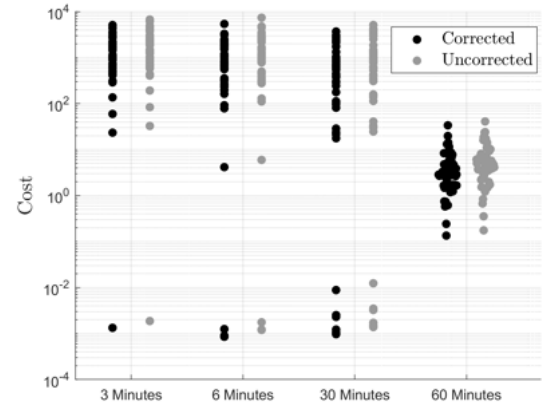

Figure 5.5d: Low Earth orbit to Geostationary cost before and after process noise correction.

flight between measurements increased to an hour. The impact of the process noise grows with time, and the linearization assumptions began to break down as well. At an hour, the results no longer clustered around 1.4 and spread out much further. The distribution was much larger for the LEO to GEO cases. However, the majority of costs shifted less than the cislunar cases. The geostationary target case has a high cost due to lower observability for the short time of flight cases. The costs don't indicate an observable results until the time of flight increased to an hour.

5.3.5 Response to Measurement Error

Any measurement of the relative angle between an observer and target will have some amount of measurement error. In order to perform orbit determination on a series of measurements of an object hundreds of thousands of kilometers away, the algorithm applied must be capable of handling some level of measurement error. The response of the proposed algorithm to measurement error is an important metric for assessing its capabilities under realistic conditions. Simulated scenarios are generated with a high-fidelity orbit propagator for several orbital regimes.

Each regime has two hundred cases randomly generated within the defining bounds of the regime. Due to the large distances involved in orbit determination for cislunar space, the scenarios are generated with two different amounts of measurement error standard deviations, 0.1 arcseconds and 1 arcsecond. The errors are Gaussian distributed about the truth, and the error ellipse is a circle projected onto the unit sphere. The time delay between each simulated observation is six hours. The Nelder-Mead algorithm is set to converge to a tolerance between the minimum vertex cost and the maximum vertex cost of 0.01, and the Nelder-Mead algorithm is set to reinitialize up to 10 times or until a result has a minimal cost below 1. This is a much tighter constraint than the developed necessary condition for convergence to a potential solution, and considering that the cost is a average of Mahalanobis distances, a cost below 1 is indistinguishable from the truth. This is because if the average of Mahalanobis distances is less than one, then the projections are well within the expected range for an accurate estimate. An error of only one standard deviation could reasonably be expected for an accurate estimate. This constraint allows the algorithm to exit early when a likely solution is found. If no run of the Nelder-Mead algorithm yields a result with cost below one, the lowest cost solution is returned.

The cumulative distribution functions for the estimate errors are shown in figures (5.6) through (5.9). Figures (5.6) and (5.7) show the 0.1 arcsecond measurement error case for both position and velocity initial state estimate errors respectively while figures (5.8) and (5.9) show the position and velocity errors for the 1 arcsecond case. The five plotted regimes are for a geostation-

ary observer to a L1 halo in and out of plane orbit and a Lunar Gateway inspired near rectilinear halo orbit. As well as L1 halo in plane to L2 halo out of plane and L1 halo out of plane to L2 halo in plane.

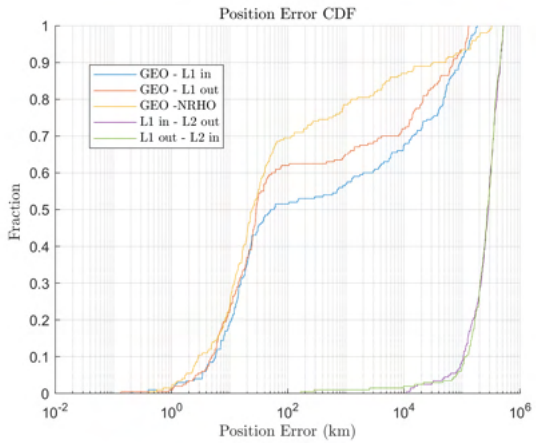

Figure 5.6: Cumulative distribution function of the position errors under 0.1 arcsecond measurement error, 10 max reinitializations, and a convergence tolerance of 0.01.

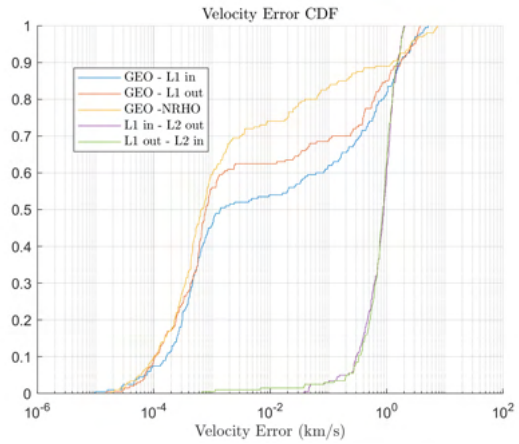

Figure 5.7: Cumulative distribution function of the velocity errors under 0.1 arcsecond measurement error, 10 max reinitializations, and a convergence tolerance of 0.01.

With a measurement error of 0.1 arcseconds, 52% of GEO to L1 halo in plane, 63% of GEO to L1 halo out of plane, and 70% of GEO to NRHO cases returned estimates within 100 kilometers of truth and after this point there is a clear point of inflection where the results transition from near estimates of the truth and cases where it never approached the truth. The two regimes observing L2 points performed poorly with only one near estimate to the truth. These results are mirrored nearly exactly before the point of inflection in the velocity distribution. This is because these accurate position estimates tended to also have accurate velocity estimates as the full state is estimated.

For the case with measurement angle standard deviation of 1 arc second, figures 5.8 and 5.9 show a reduced capability to reliably converge to the solution. The GEO to L1 in plane case has 29% below 100 kilometers while GEO to L1 out of plane and GEO to NRHO are at 36% and 39% respectively. The L2 cases remain inaccurate, and the general behavior follows the same overall trends. The algorithm shows an expected response to measurement error where measurement error is inversely correlated with estimate accuracy.

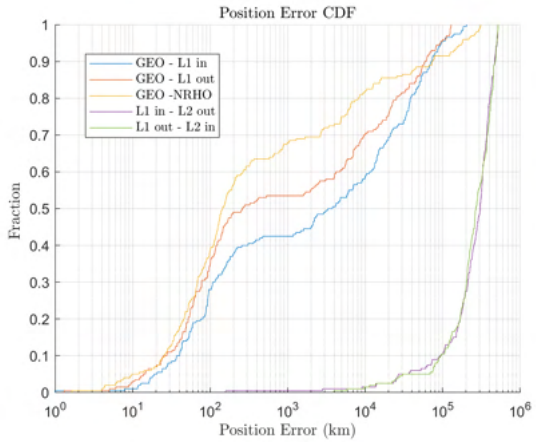

Figure 5.8: Cumulative distribution function of the position errors under 1 arcsecond measurement error, 10 max reinitializations, and a convergence tolerance of 0.01.

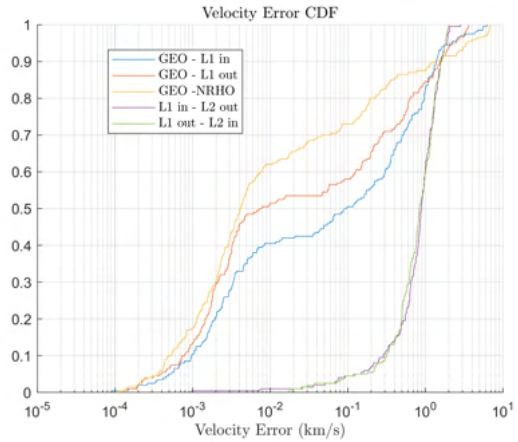

Figure 5.9: Cumulative distribution function of the velocity errors under 1 arcsecond measurement error, 10 max reinitializations, and a convergence tolerance of 0.01.

5.3.6 Impact of Optimizer Parameters

The Nelder-Mead optimizer has several customizable parameters that impact the algorithm's performance in both speed and accuracy. Since multiple minima are expected in the cost function, the algorithm is also run multiple times for each case so more of the cost function can be explored in case early runs fall into local minima. Both the number of reinitialization and the convergence tolerance for determining when to exit a run are explored for their impact on the algorithm's accuracy. From the contour plot study, it is known that in some orbital scenarios the main cost valley is long and narrow, so a convergence tolerance that is too high can lead to exiting a run too early while a tolerance too low can waste time when a solution is already found. The investigated parameters are the convergence tolerance and the number of reinitializations. Since the appropriate balance between time cost and accuracy depends on the implementation goals, these results are intended to show the trade-offs in choosing these values rather than provide a single, universal solution.

Figures (5.10) and (5.11) show cumulative distribution functions for the position errors for cases with five and twenty reinitializations respectively. Contrast these with figure (5.6) which

shows the results for ten reinitializations. The case with five reinitializations shows a sharp drop-off in the algorithm's accuracy. The GEO to L1 halo in plane regime has 31% of solutions within one hundred kilometers. The GEO to L1 halo out of plane and GEO to NRHO have 42% and 58% respectively. By contrast, the case with twenty reinitializations shows an improvement in accuracy and a convergence of the three cases observed from GEO. The percent of results under one hundred kilometers from the true initial position for GEO to L1 halo in plane, GEO to L1 halo out of plane, and GEO to NRHO are 73%, 78%, and 75% respectively. The cases observing L2 orbits from L1 orbits remain inaccurate.

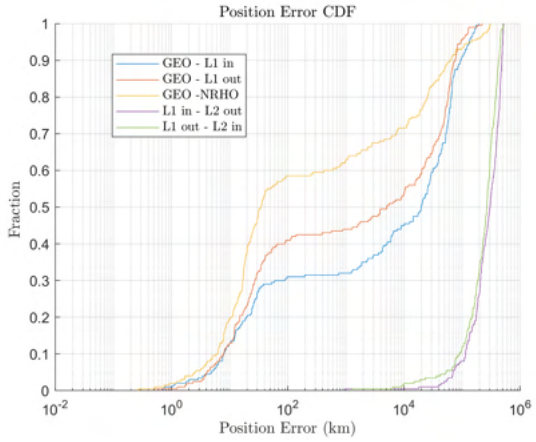

Figure 5.10: Cumulative distribution function of the position errors under 0.1 arcsecond measurement error, 5 max reinitializations, and a convergence tolerance of 0.01.

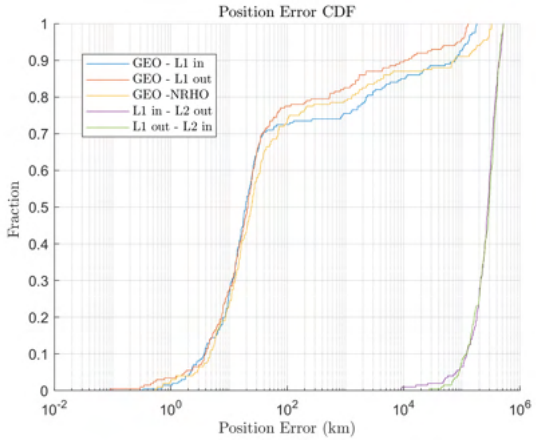

Figure 5.11: Cumulative distribution function of the position errors under 0.1 arcsecond measurement error, 20 max reinitializations, and a convergence tolerance of 0.01.

The other investigated optimizer parameter is the convergence tolerance. This is implemented as a difference between the maximal and minimal cost of a simplex's vertices. When this falls below the given tolerance, the Nelder-Mead optimzer exits the run and keeps the minimal vertex as the solution. Tolerances of 0.1, 0.01, and 0.001 are investigated. The results for the case with a tolerance of 0.01 are shown in figure (5.6) while the results for tolerances of 0.1 and 0.001 are shown in figures (5.12) and (5.13) respectively. A smaller tolerance will require more iterations to converge, but may converge to a value closer to the truth. This is shown in these results. When the tolerance is 0.1, the point of inflection in the plot moves toward higher position error states. This is

due to a subset of Nelder-Mead runs falling into the appropriate cost well, but exiting before they reach the bottom due to the looser constraint. Similarly, with the tighter constraint of 0.001, the point of inflection moves toward smaller position errors. It becomes just above ten kilometers in error rather than a bit above one hundred kilometers as in the case with a tolerance of 0.01. The portion of near-truth results also increases for the tighter cost case and decreases for the looser cost case.

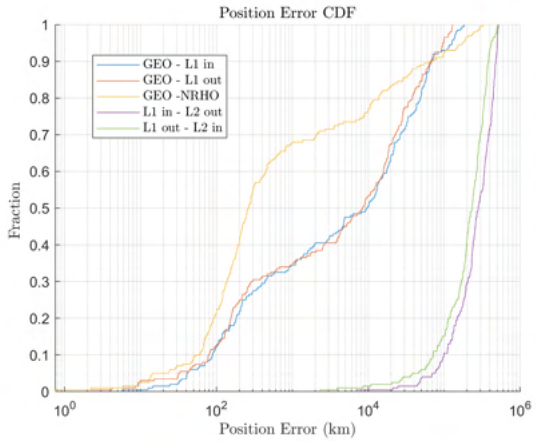

Figure 5.12: Cumulative distribution function of the position errors under 0.1 arcsecond measurement error, 10 max reinitializations, and a convergence tolerance of 0.1.

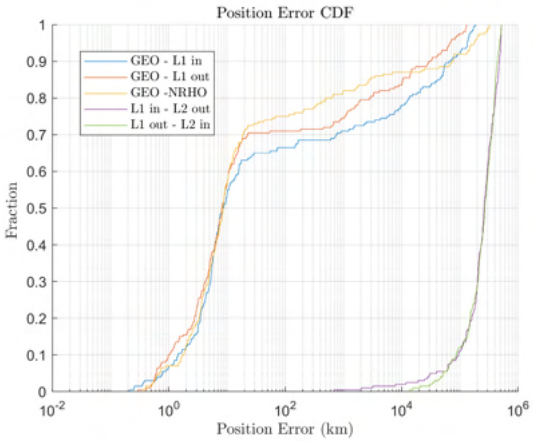

Figure 5.13: Cumulative distribution function of the position errors under 0.1 arcsecond measurement error, 10 max reinitializations, and a convergence tolerance of 0.001.

In all assessed cases, the scenarios with observers in L1 orbits and targets in L2 orbits fail to converge to accurate state estimates. This issue persists even if the number of reinitializations is increased to fifty and the full set of solutions is saved. Figure (5.14) shows the results for such a case with the closest to truth solution from the returned set of fifty kept as the case's solution. However, despite the failure to return a set of solutions containing the truth, when the algorithm is manually initialized near the truth it converges to it. The dynamics of this regime complicate the search. Orbits about Lagrange points one and two are unstable, and small differences in the dynamics for the truth and estimate can lead to divergent results as time increases. In addition, the solution space has several minima deep enough to appear indistinguishable from the truth, and the region around the true state in the solution space is small enough that even with fifty runs, the

algorithm is more likely to miss the true state than to capture it. Since the other solutions that are discovered have sufficiently low cost, these states correspond to other orbits that could feasibly produce similar measurements. Accordingly, eliminating these solutions would require ruling out the corresponding states, despite being otherwise feasible. Without a priori knowledge that can restrict the search space further, these other solutions keep the results of a search in this regime ambiguous.

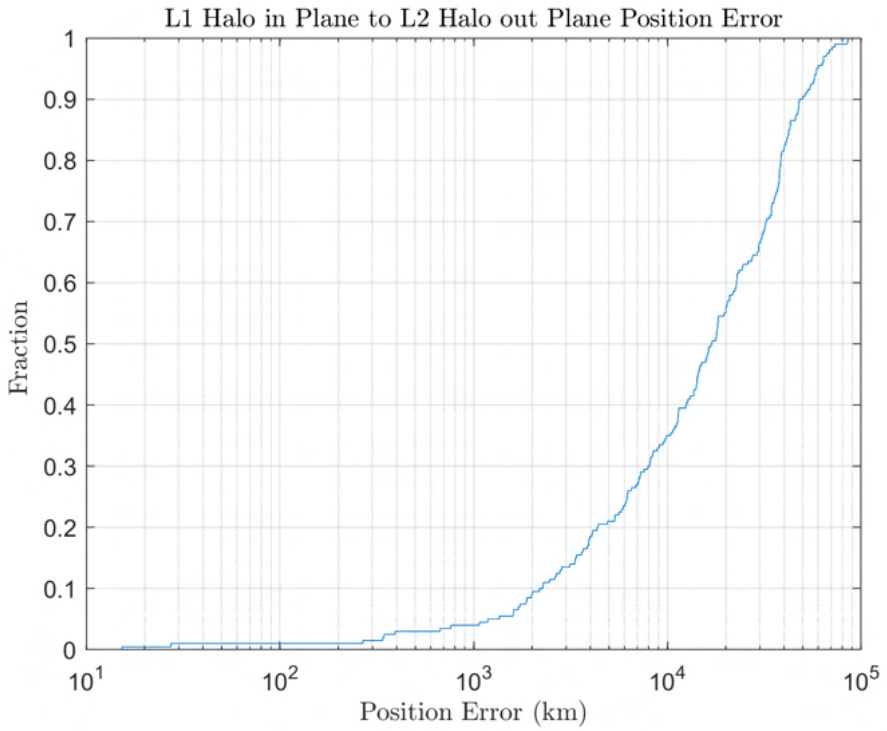

Figure 5.14: CDF of the best solution over 50 reinitializations for 200 cases of L1 halo in plane observers to L2 halo out of plane targets.

5.4 Chapter Summary

An optimization-based algorithm for performing initial orbit determination under arbitrary dynamics has been developed and it's accuracy assessed under simulated scenarios. In addition, an analytical method for determining if a proposed solution satisfies necessary but not sufficient conditions to be a solution for a given set of measurements with consideration for process noise

has been developed. To explore the cost-benefit ratio of several parameters in the dynamics, the time-cost and propagated state accuracy were assessed. This was done by propagated the orbits of randomly generated objects in various regimes under several simplifications to the dynamics. The results were compared against a high-fidelity propagation and it was discovered that a second order geopotential model for the Earth and Moon with the dynamics of all major Solar System bodies included represented a good balance between propagation accuracy and computational cost.

Contours of the cost function of the algorithm were then developed for a few interesting cases, and their geometry was explored for the impact that they may have on the optimization. It was discovered that the contour had several local minima which needed to be accounted for the algorithm implementation. The algorithm was then run over randomly generated scenarios within a set list of orbital regimes. The behavior of the algorithm in response to adjustments to optimizer parameters as well as measurement error was explored yielding a series of plots showing the relative success of the algorithm subject to these adjusted scenarios. With low measurement error and sufficiently generous optimizer parameters, the proposed IOD algorithm can return a close estimate to the truth as the best fit over 70% of the time.

Chapter 6

Conclusions and Future Work

6.1 Conclusion

In the previous few chapters four major contributions were discussed for advancements in optical space domain awareness in the measurement to state estimate pipeline. This started with improvements to astrometric measurement in ultra-wide angle images. Then it moved into methods for extracting photometric light curves from streaks. Finally methods for initial orbit determination in several domains were developed and discussed. These methods can be used separately or together in order to extract states from observations with a focus on information optimality and a robust response to a variety of scenarios. Improvements from earlier stages in the process can feed into later steps so that the final estimate is more accurate thanks to each contribution should each step be applicable and applied.

In the first contribution, empirical methods for calibrating ultra-wide angle lenses were developed. These methods were capable of estimating and compensating for distortion in images that would not be well-modeled by existing function-based methods. These approaches were tested against both simulated and observed data from the OmniSSA system. The calibration residuals across the frame were collected to compare errors both before and after calibration.

Image calibration is an essential step for extracting meaningful data from electro-optical observations of space objects. It allows for the accurate measurement of the relative position of observed phenomena. The relative angles and angle rates to an object can be used to build a full dynamical state estimate. In addition to the astrometric measurement of an object, the photometric

information can be used to build a picture of the rotational state or build a deeper understanding of the translational state [15]. To extract as much useful information from an observed streaking object as possible, both the calibrated position measurements as well as the photometric light curve of the object should be measured. With the astrometic measurements improved from the image calibration, the photometric information in an observed object can be more reliably extracted in order to find the object's light curve.

Streaking objects such as those discussed in the second contribution encode information about their rotational and translational state. Photometric information can be used to determine if an object is entering the Earth's penumbra, and finding the precise length of the streak can reveal the object's relative angular velocity. With two images containing both angle and angle rate information, a full state solution can potentially be formed.

A pair of methods for photometric light curve extraction were developed in the this chapter. These approaches offered a way to determine the optimal number of streak segments to break a streak into in order to extract all available information without attempting to estimate an unobservable quantity of streak segments. The photometric light curve could be used to estimate a target object's attitude state or if it had potentially passed into the Earth's penumbra. These methods were assessed against simulated and observed data to show that the two methods provided consistent results and that they extracted all available photometric information in the streak. In the next chapter, an initial orbit determination method is developed to take this kind of information and build a full state solution.

The robust method for near-Earth IOD followed. A more robust approach to initial orbit determination allows an operator to apply a single method to incoming data without requiring any a priori knowledge of the target's state. This simplifies the process of finding state estimates and reduces the potential for confusion when multiple algorithms produce different results with the same input data. The IOD algorithm described in chapter four trades some computational efficiency for reliability under a variety of scenarios. This makes the algorithm broadly applicable for observations with measurements composed of both angles and angle rates.

This optimization-based method was shown to work well in a variety of orbital scenarios without requiring an initial guess. It was verified against both simulated and observed data and compared against other popular methods for IOD. One of its advantages is that it works with both angles and angle rates so data derived from streaking objects can be easily applied as input. The final major contribution gave up the two-body assumption and some of its advantages in computation time and accuracy to extend this method to beyond the near-Earth environment.

The work of the previous chapter was then adapted with significant changes to work beyond the near-Earth environment. Considering the lack of reliable methods for initial orbit determination beyond the near-Earth environment, an approach that could work without the two body assumption could enable more reliable orbit determination in a domain currently seeing rapidly growing interest. As orbits move away from the Earth, the influence of other factors become more prominent and the assumptions of traditional IOD fall apart. The optimization based approach of the previous chapter was then adjusted to propagate between orbits using a full dynamics simulation so that it could accept arbitrary dynamics to model the increasingly complex space.

Extending the algorithm to orbits beyond the near-Earth domain required significant changes to the definition of the cost function in both the way it was derived and a transition to a more meaningful cost value. Some of these changes improved results without individually requiring the transition away from the two-body formulation. These could then be back ported into the two body algorithm to provide more meaningful values there. This includes the inclusion of process noise in the cost so that the Mahalanobis distance closer represented the actual uncertainty of the estimate.

6.1.1 Impact on the Literature

Prior to the work presented here, the state of the art as described by the literature lacked formal information-optimal methods for extracting photometric light curves from streaks as well as approaches for initial orbit determination beyond the near-Earth save for implementations of the batch least squares filter which struggled to converge when there were multiple potential solutions.

The work relating to light curve extraction here provides a new formal process for optimally extracting this information from images and reliably conveying the uncertainty in the resulting light curve estimate. Similarly, the final contribution provides a process to explore the solution space of the cislunar and beyond IOD problem. This allows the set of potential solutions to be found for these problems which is an important step in building accurate estimates of the dynamical state of objects in these domains. In addition, the resulting cost characterizes how well the estimate fits the measurements in a Mahalanobis distance that relates the error to the expected error based on the measurement covariance and process noise.

There exist methods for calibrating images to compensate for distortion. However, these use the function-based approach that necessarily assume some symmetry in the distortion. The empirical approach described in this research provides a method for calibrating images without assuming any symmetry. This provides new utility that is otherwise nonexistent in the current literature. Similar to the image calibration research, there are many existing methods for near-Earth initial orbit determination. However, the method presented here is much more robust to many different scenario parameters including the relative orbits of both the observer and target as well as the time between observations. In addition, most existing methods are intended for three angles-only observations. This approach offers a solution for cases with data derived from streaks that contain both angles and angle rate data.

6.1.2 Impact on Society

These contributions build an observation to estimate pipeline for optical data with improvements over the state of the art at every step. Together these methods can enable approaches to space domain awareness not previously viable in environments with new focus from the major stake holders in the space domain. The information optimal approaches implemented here can yield more rigorous results for generating estimates of a space object's states. Together this allows observers to build a better understanding of the state of the space environment and can reduce the potential for collisions and failed missions from poor data and estimates.

6.2 Future Work

In the calibration method, there remains room for improvement particularly in the method by which observed and catalog stars were associated with one another. The applied method only captures a small fraction of the visible stars in an image and was prone to misassociating stars in regions with larger residuals. The Hungarian algorithm was not robust enough to reliably correctly associate the calibrating stars. Potential more robust approaches include Lagrangian relaxation, Monte-Carlo tree search, machine learning, and approaches based on associating geometries of combinations of stars rather than minimizing distance residuals [93] [94]. If this approach for empirical calibration were combined with a robust method for star association, accuracy could approach that of the simulated cases.

For the photometric light curve extraction work, the method would theoretically also work as a useful tool for measuring spectra from slitless spectroscopy images. The proposed methods would be able to find a spectrum to the appropriate wavelength resolution for the information contained in the spectrum streak. This would need to be developed further to find what kinds of complications may arise with this type of data. Aside from the photometric information, an imaged streak provides measurements for both angles and angle rates in a single observation. Two of these would be required to perform initial orbit determination from observed streaks.

In the cislunar IOD approach, the algorithm may be able to be adjusted in order to produce a covariance matrix for the state estimate. This would include process noise which is not possible with a batch least squares approach. If I can figure out a way to compute the state covariance, the result could then be tested against a batch least squares filter. If the results are similar, and they properly optimize the same value, that would indicate that this algorithm could be used in place of batch least squares when process noise is needed in the resulting estimate. As the results of initial orbit determination are typically passed through a batch least squares filter to refine their value and find the state uncertainty, this would allow both steps to be combined into one without giving up on including process noise.

Bibliography

[1] Adam Danz. boxplotgroup. `https://www.mathworks.com/matlabcentral/fileexchange/74437-boxplotgroup`, MATLAB Central File Exchange, 2020. Retrieved May 21, 2020.

[2] John A Kennewell and Ba-Ngu Vo. An overview of space situational awareness. In Proceedings of the 16th International Conference on Information Fusion, pages 1029–1036. IEEE, 2013.

[3] Travis Blake, M Sánchez, J Krassner, M Georgen, and S Sundbeck. Space domain awareness. Technical report, DEFENSE ADVANCED RESEARCH PROJECTS AGENCY ARLINGTON VA, 2012.

[4] Brian Weeden, Paul Cefola, and Jaganath Sankaran. Global space situational awareness sensors. In AMOS Conference, 2010.

[5] Laurence G Taff. On initial orbit determination. The Astronomical Journal, 89:1426–1428, 1984.

[6] Takayuki Ito and Yasuyuki Sugano. Fast ultra-wide angle lens system, December 19 1995. US Patent 5,477,389.

[7] Ryan D Coder and Marcus J Holzinger. Multi-objective design of optical systems for space situational awareness. Acta Astronautica, 128:669–684, 2016.

[8] Željko Ivezić, Steven M Kahn, J Anthony Tyson, Bob Abel, Emily Acosta, Robyn Allsman, David Alonso, Yusra AlSayyad, Scott F Anderson, John Andrew, et al. Lsst: from science drivers to reference design and anticipated data products. The Astrophysical Journal, 873(2):111, 2019.

[9] Roberto G Abraham and Pieter G van Dokkum. Ultra–low surface brightness imaging with the dragonfly telephoto array. Publications of the Astronomical Society of the Pacific, 126(935):55, 2014.

[10] Janne Heikkila, Olli Silven, et al. A four-step camera calibration procedure with implicit image correction. In cvpr, volume 97, page 1106. Citeseer, 1997.

[11] Mark R Ackermann, David D Cox, John McGraw, and Peter Zimmer. Lens and camera arrays for sky surveys and space surveillance. Technical report, Sandia National Lab.(SNL-NM), Albuquerque, NM (United States), 2016.

[12] Duane C Brown. Decentering distortion of lenses. Photogrammetric Engineering and Remote Sensing, 1966.

[13] Andrew W Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, volume 1, pages I–I. IEEE, 2001.

[14] Haider Taha. The potential for air-temperature impact from large-scale deployment of solar photovoltaic arrays in urban areas. Solar Energy, 91:358–367, 2013.

[15] Richard Linares, Moriba K Jah, John L Crassidis, and Christopher K Nebelecky. Space object shape characterization and tracking using light curve and angles data. Journal of Guidance, Control, and Dynamics, 37(1):13–25, 2014.

[16] Jiri Silha, Esther Linder, Monika Hager, and Thomas Schildknecht. Optical light curve observations to determine attitude states of space debris. In Proceedings of 30th International Symposium on Space Technology and Science, Kobe-Hyogo, Japan, 2015.

[17] Paul Chote, James A Blake, and Don Pollacco. Precision optical light curves of leo and geo objects. 2019.

[18] David A Vallado. Fundamentals of astrodynamics and applications, volume 12. Springer Science & Business Media, 2001.

[19] Reza Raymond Karimi and Daniele Mortari. A performance based comparison of angle-only initial orbit determination methods. Adv. Astronaut. Sci., AAS/AIAA, Hilton Head Island, South Carolina, 150:1793–1809, 2013.

[20] Pedro Ramon Escobal. Methods of orbit determination. New York: Wiley, 1965, 1965.

[21] RH Gooding. A procedure for the solution of lambert's orbital boundary-value problem. Celestial Mechanics and Dynamical Astronomy, 48(2):145–165, 1990.

[22] Greg Chavers, Nantel Suzuki, Marshall Smith, Lisa Watson-Morgan, Steven W Clarke, Walter C Engelund, Lindsay Aitchison, Shawn McEniry, Laura Means, Michael DeKlotz, et al. Nasa's human lunar landing strategy. 2019.

[23] Timothy S Murphy and Marcus J Holzinger. Uncued low snr detection with likelihood from image multi bernoulli filter. In Advanced Maui Optical and Space Surveillance Technologies Conference, 2016.

[24] Shishir Shah and JK Aggarwal. Intrinsic parameter calibration procedure for a (high-distortion) fish-eye lens camera with distortion model and accuracy estimation. Pattern Recognition, 29(11):1775–1788, 1996.

[25] Jianhua Wang, Fanhuai Shi, Jing Zhang, and Yuncai Liu. A new calibration model of camera lens distortion. Pattern Recognition, 41(2):607–615, 2008.

[26] Gideon P Stein. Lens distortion calibration using point correspondences. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 602–608. IEEE, 1997.

[27] Saijiang Ai, Xiang Wang, Mengchao Ma, and Keyi Wang. A method for correcting non-linear geometric distortion in ultra-wide-angle imaging system. Optik, 124(24):7014–7021, 2013.

[28] Jenni Virtanen, Jonne Poikonen, Tero Säntti, Tuomo Komulainen, Johanna Torppa, Mikael Granvik, Karri Muinonen, Hanna Pentikäinen, Julia Martikainen, Jyri Näränen, et al. Streak detection and analysis pipeline for space-debris optical images. Advances in Space Research, 57(8):1607–1623, 2016.

[29] Hirohisa Kurosaki, Toshifumi Yanagisawa, and Atsushi Nakajima. Observation of light curves of space objects. In AMOS conference, 2009.

[30] Reza Raymond Karimi and Daniele Mortari. Initial orbit determination using multiple observations. Celestial Mechanics and Dynamical Astronomy, 109(2):167–180, 2011.

[31] Reza Raymond Karimi and Daniele Mortari. Orbit determination using prescribed orbits. In AAS/AIAA Space Flight Mechanics Meeting Conference, pages 2010–236, 2010.

[32] Kyle J DeMars and Moriba K Jah. Probabilistic initial orbit determination using gaussian mixture models. Journal of Guidance, Control, and Dynamics, 36(5):1324–1335, 2013.

[33] LG Taff and DL Hall. The use of angles and angular rates. Celestial mechanics, 16(4):481–488, 1977.

[34] Laurence G Taff. The resurrection of laplace's method of initial orbit determination. Technical report, MASSACHUSETTS INST OF TECH LEXINGTON LINCOLN LAB, 1983.

[35] Carl Friedrich Gauss. Theoria motus corporum coelestium in sectionibus conicis solem ambientium, volume 7. Perthes et Besser, 1809.

[36] Anread Milani, Giovanni 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, 2004.

[37] Giacomo Tommei, Andrea Milani, and Alessandro Rossi. Orbit determination of space debris: admissible regions. Celestia Mechanics and Dynamical Astronomy, 97(4):289–304, 2007.

[38] 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, 2009.

[39] Davide Farnocchia, Giacomo Tommei, Andrea Milani, and Alessandro Rossi. Innovative methods of correlation and orbit determination for space debris. Celestial Mechanics and Dynamical Astronomy, 107(1-2):169–185, 2010.

[40] Marcus J Holzinger. Initial orbit determination, data association, and admissible regions of space objects using magnetometers. Journal of Guidance, Control, and Dynamics (accepted in April, 2014), 2014.

[41] Samuel Wishnek and Marcus J Holzinger. Astrometry and time-resolved photometry from streak using calibrated ultra-wide field of view cameras. In AIAA Scitech 2020 Forum, page 1592, 2020.

[42] Samuel Wishnek, Marcus J Holzinger, and Patrick Handley. Robust initial orbit determination using streaks and admissible regions. In AMOS 2020, 2020.

[43] Samuel Wishnek and Marcus J Holzinger. Empirical approaches to ultra-wide angle lens calibration. In 31ST AAS/AIAA Space Flight Mechanics Meeting, 2021.

[44] Samuel Wishnek, Marcus J Holzinger, and Patrick Handley. Robust cislunar initial orbit determination. In AMOS 2021, 2021.

[45] Sam Wishnek, Marcus J Holzinger, Patrick Handley, and Sue Hagerty. Robust initial orbit determination using streaks and admissible regions. The Journal of the Astronautical Sciences, pages 1–42, 2021.

[46] Samuel Wishnek and Marcus J Holzinger. Empirical approaches to ultra-wide angle lens calibration. Journal of Astronautical Sciences (submitted), 2021.

[47] Samuel Wishnek and Marcus J Holzinger. Optimal extraction of photometric light curves from space object streaks. Journal of Astronautical Sciences (submitted), 2021.

[48] Jason P De Villiers, F Wilhelm Leuschner, and Ronelle Geldenhuys. Centi-pixel accurate real-time inverse distortion correction. In Optomechatronic Technologies 2008, volume 7266, page 726611. International Society for Optics and Photonics, 2008.

[49] Lale Özbakir, Adil Baykasoğlu, and Pınar Tapkan. Bees algorithm for generalized assignment problem. Applied Mathematics and Computation, 215(11):3782–3795, 2010.

[50] Boris Delaunay et al. Sur la sphere vide. Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk, 7(793-800):1–2, 1934.

[51] William H Press and Saul A Teukolsky. Savitzky-golay smoothing filters. Computers in Physics, 4(6):669–672, 1990.

[52] Chunsheng Yu and Qingjin Peng. Robust recognition of checkerboard pattern for camera calibration. Optical Engineering, 45(9):093201, 2006.

[53] Zhongshi Wang, Zhongge Wang, and Xinhe Xu. Extraction of the corner of checkerboard image. In 2008 7th World Congress on Intelligent Control and Automation, pages 6789–6792. IEEE, 2008.

[54] ESA. The hipparcos and tycho catalogues. ESA SP-1200, 1997.

[55] Dustin Lang, David W Hogg, Keir Mierle, Michael Blanton, and Sam Roweis. Astrometry. net: Blind astrometric calibration of arbitrary astronomical images. The astronomical journal, 139(5):1782, 2010.

[56] Harold W Kuhn. The hungarian method for the assignment problem. Naval research logistics quarterly, 2(1-2):83–97, 1955.

[57] David F Watson. Computing the n-dimensional delaunay tessellation with application to voronoi polytopes. The computer journal, 24(2):167–172, 1981.

[58] Enobong Bassey, Jacqueline Whalley, and Philip Sallis. An evaluation of smoothing filters for gas sensor signal cleaning. In Proceedings of the Fourth International Conference on Advanced Communications and Computation, Paris, France, pages 20–24, 2014.

[59] Hai-Bin Pan, Wei Zhang, and Ming-Yu Cong. Detection algorithm for space dim moving object. In Fundamental Problems of Optoelectronics and Microelectronics III, volume 6595, page 65951H. International Society for Optics and Photonics, 2007.

[60] J Sharma, C Von Braun, and EM Gaposchkin. Space-based visible data reduction. Journal of Guidance, Control, and Dynamics, 23(1):170–174, 2000.

[61] D O'Donoghue. High speed ccd photometry. Baltic Astronomy, 4:519–526, 1995.

[62] Julien Demouth, Olivier Devillers, Hazel Everett, Marc Glisse, Sylvain Lazard, and Raimund Seidel. Between umbra and penumbra. In Proceedings of the twenty-third annual symposium on Computational geometry, pages 265–274, 2007.

[63] Thomas Schildknecht, Reto Musci, Carolin Früh, and Martin Ploner. Color photometry and light curve observations of space debris in geo. In Proceedings of Advanced Maui Optical and Space Surveillance Technologies Conference, pages 17–19, 2008.

[64] John W Hardy, J E_ Lefebvre, and CL Koliopoulos. Real-time atmospheric compensation. JOSA, 67(3):360–369, 1977.

[65] Zuo Wang, PengFeng Xiao, XingFa Gu, XueZhi Feng, XiaoYing Li, HaiLiang Gao, Hui Li, JinTang Lin, and XueLiang Zhang. Uncertainty analysis of cross-calibration for hj-1 ccd camera. Science China Technological Sciences, 56(3):713–723, 2013.

[66] WJ Merline and Steve B Howell. A realistic model for point-sources imaged on array detectors: The model and initial results. Experimental Astronomy, 6(1-2):163–210, 1995.

[67] S Ya Shatskikh. Multivariate cauchy distributions as locally gaussian distributions. Journal of Mathematical Sciences, 78(1):102–108, 1996.

[68] Dan Simon. Optimal state estimation: Kalman, H infinity, and nonlinear approaches. John Wiley & Sons, 2006.

[69] P Sebastiani and Henry P Wynn. Experimental design to maximize information. In AIP Conference Proceedings, volume 568, pages 192–203. AIP, 2001.

[70] Paul R Rider. Variance of the median of samples from a cauchy distribution. Journal of the American Statistical Association, 55(290):322–323, 1960.

[71] Moawwad EA El-Mikkawy. On the inverse of a general tridiagonal matrix. Applied Mathematics and Computation, 150(3):669–679, 2004.

[72] Martin Kümmel, JR Walsh, Norbert Pirzkal, Harald Kuntschner, and Anna Pasquali. The slitless spectroscopy data extraction software axe. Publications of the Astronomical Society of the Pacific, 121(875):59, 2009.

[73] Bob Schutz, Byron Tapley, and George H Born. Statistical orbit determination. Elsevier, 2004.

[74] John E Prussing and Bruce A Conway. Orbital mechanics. Oxford University Press, USA, 1993.

[75] Bruce A Conway. An improved algorithm due to laguerre for the solution of kepler's equation. Celestial mechanics, 39(2):199–211, 1986.

[76] Yuhui Shi and Russell C Eberhart. Parameter selection in particle swarm optimization. In International conference on evolutionary programming, pages 591–600. Springer, 1998.

[77] John A Nelder and Roger Mead. A simplex method for function minimization. The computer journal, 7(4):308–313, 1965.

[78] Marco A Luersen and Rodolphe Le Riche. Globalized nelder–mead method for engineering optimization. Computers & structures, 82(23-26):2251–2260, 2004.

[79] Jeffrey C Lagarias, James A Reeds, Margaret H Wright, and Paul E Wright. Convergence properties of the nelder–mead simplex method in low dimensions. SIAM Journal on optimization, 9(1):112–147, 1998.

[80] Andrew Vernon Schaeperkoetter. A comprehensive comparison between angles-only initial orbit determination techniques. PhD thesis, Texas A & M University, 2012.

[81] Space track. `https://www.space-track.org/`, 5 2020.

[82] Stellarium. `http://stellarium.sourceforge.net/`, 5 2020. Version 0.16.1.

[83] David A Vallado and Paul J Cefola. Two-line element sets–practice and use. In 63rd International Astronautical Congress, Naples, Italy, 2012.

[84] Charles H Acton Jr. Ancillary data services of nasa's navigation and ancillary information facility. Planetary and Space Science, 44(1):65–70, 1996.

[85] Nikolaos K Pavlis, Simon A Holmes, Steve C Kenyon, and John K Factor. The development and evaluation of the earth gravitational model 2008 (egm2008). Journal of geophysical research: solid earth, 117(B4), 2012.

[86] S Goossens, FG Lemoine, TJ Sabaka, JB Nicholas, E Mazarico, DD Rowlands, BD Loomis, DS Chinn, GA Neumann, DE Smith, et al. A global degree and order 1200 model of the lunar gravity field using grail mission data. In Lunar and Planetary Science Conference, number 1903, page 1484, 2016.

[87] Johnny L Worthy III, Marcus J Holzinger, and Kohei Fujimoto. Optical sensor constraints on space object detection and admissible regions. Paper AAS, pages 13–707, 2013.

[88] Johnny L Worthy III, Marcus J Holzinger, and Daniel J Scheeres. An optimization approach for observation association with systemic uncertainty applied to electro-optical systems. Advances in Space Research, 61(11):2709–2724, 2018.

[89] Lixing Han and Michael Neumann. Effect of dimensionality on the nelder–mead simplex method. Optimization Methods and Software, 21(1):1–16, 2006.

[90] John Michael Tutill Thompson and H Bruce Stewart. Nonlinear dynamics and chaos. John Wiley & Sons, 2002.

[91] Jing Ren, Mingtao Li, and Jianhua Zheng. Families of transfers from the moon to distant retrograde orbits in cislunar space. Astrophysics and Space Science, 365(12):1–21, 2020.

[92] Marcus J Holzinger, Daniel J Scheeres, and Kyle T Alfriend. Object correlation, maneuver detection, and characterization using control distance metrics. Journal of Guidance, Control, and Dynamics, 35(4):1312–1325, 2012.

[93] 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(2):129–150, 1997.

[94] Cameron B Browne, Edward Powley, Daniel Whitehouse, Simon M Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. A survey of monte carlo tree search methods. IEEE Transactions on Computational Intelligence and AI in games, 4(1):1–43, 2012.

[95] Evan Kaufman, T Alan Lovell, and Taeyoung Lee. Nonlinear observability for relative orbit determination with angles-only measurements. The Journal of the Astronautical Sciences, 63(1):60–80, 2016.

[96] M Levesque. Automatic reacquisition of satellite positions by detecting their expected streaks in astronomical images. In Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference, page E81, 2009.

[97] Daniel Aguilar Marsillach, Shahzad Virani, and Marcus J Holzinger. Iaa-icssa-17-0x-xx real-time hardware-in-the-loop hand-off from a finder scope to a larger telescope. ICSSA, 2017.

[98] Felipe I San Martín, Claudio A Perez, Juan E Tapia, Shahzad Virani, and Marcus J Holzinger. Automatic space object detection on all-sky images from a synoptic survey synthetic telescope array. Advances in Space Research, 2019.

[99] AD Barbour. Poisson convergence and random graphs. In Mathematical Proceedings of the Cambridge Philosophical Society, volume 92, pages 349–359. Cambridge University Press, 1982.

[100] Roger R Bate, Donald D Mueller, Jerry E White, and William W Saylor. Fundamentals of astrodynamics. Courier Dover Publications, 2020.

[101] Giorgio Mingotti, Francesco Topputo, and Franco Bernelli-Zazzera. Efficient invariant-manifold, low-thrust planar trajectories to the moon. Communications in Nonlinear Science and Numerical Simulation, 17(2):817–831, 2012.

[102] William Oberkampf, Jon Helton, and Kari Sentz. Mathematical representation of uncertainty. In 19th AIAA applied aerodynamics conference, page 1645, 2001.

[103] Marcus J Holzinger. Dynamic data driven electro-optical sensor detection, tracking, and multi-objective control. Air Force Fiscal Year 2017 Young Investigator Research Program, June 2016.

[104] Abraham Savitzky and Marcel JE Golay. Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry, 36(8):1627–1639, 1964.