The Report Committee for Alexander Chiu
certifies that this is the approved version of the following report:

Admissible Region Splitting for Multi-Fidelity Orbit Propagation

SUPERVISING COMMITTEE:

Brandon A. Jones, Supervisor

Moriba K. Jah

Admissible Region Splitting for Multi-Fidelity Orbit Propagation

by

Alexander Chiu

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

May 2023

Epigraph

Sometimes a feeling is all we humans have to go on.

– Cpt. J. T. Kirk

Acknowledgments

I am grateful for my advisor, Dr. Jones, for all his guidance and support throughout this project. Many thanks to Dr. Jah for providing insightful feedback. I am also thankful for the advice and input of Trevor, Ben, and Enrico, and the encouragement and reassurance from my friends and family. This material is based upon work supported by the Air Force Research Laboratory under agreement number FA8750-19-1-1001 and the Air Force Office of Scientific Research under award number FA9550-19-1-0404. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Air Force, the Air Force Research Laboratory, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

Abstract

Admissible Region Splitting for Multi-Fidelity Orbit Propagation

Alexander Chiu, MSE
The University of Texas at Austin, 2023

SUPERVISOR: Brandon A. Jones

Admissible regions allow for active tasking of sensors for orbit determination of newly detected objects. These regions are defined using physics-based constraint hypotheses applied to detected objects. The regions can be represented probabilistically using Gaussian mixtures and propagated forward in time, which is computationally expensive. Multi-fidelity orbit uncertainty propagation can help speed up this process, but performance degrades when applied to large initial distributions. A new method for splitting admissible regions is proposed based on k-means clustering. A sensitivity analysis is performed using the partial derivatives of orbital elements with respect to spherical coordinates. The clustering algorithm leads to enhanced multi-fidelity performance by keeping the points closer together over time. The accuracy and runtime improvements are demonstrated using simulated optical and radar test cases.

Table of Contents

Chapter 1: Introduction

The number of spacecraft in Earth orbit has continued to grow rapidly in recent years, leading to an increased need for space domain awareness and efforts to monitor anthropogenic space objects. Various organizations have created space object catalogs to track space debris and identify possible risks to ongoing missions. These catalogs typically contain orbit state estimates for different objects, and include some measure of uncertainty that reflects sensor errors and imperfectly modeled orbital dynamics. These limitations also mean that these catalogs do not contain information about all objects in orbit. As such, operators must have the capability to process observations for newly detected objects.

When a ground- or space-based sensor makes a new observation, algorithms try to associate and match it with a cataloged space object. If a data association algorithm determines that the observation fits a documented object, it can be used as a measurement update for that object. If nothing in the catalog matches the measurement, however, the object is deemed an "uncorrelated track" (Kalur et al., 2020). Operators seek to perform a full orbit determination for this object such that follow-up tracking can be performed for it. This is easier said than done, as without additional information, data from a single pass is often referred to as a "too short arc," or TSA, as a short arc of observations fails to fully characterize an orbit (Tommei et al., 2007). Instead, multiple measurements with sufficient diversity have to be fused to determine an object's orbit with confidence sufficient for tracking and prediction.

A brute force approach to this problem takes a large set of observation arcs and examines combinations of them until feasible orbits are discovered. These techniques can be effective, albeit costly, and significant research has been performed on associating different observations together. However, this process requires data over

an undesirable amount of time for risk assessment and requires that unidentified objects be detected multiple times with robust and accurate observation-to-observation association. Alternative techniques designed to mitigate these issues exist, but require increased runtime or objectionable constraints on the solution. This report presents a multi-fidelity approach to rapidly predict a single-track initial orbit determination solution based on admissible regions.

Admissible regions are able to constrain a detected object's state space using hypotheses about system dynamics. This allows operators to actively task sensors to obtain additional detections needed for orbit determination. Multi-fidelity approaches are able to greatly enhance the computation speed of this process, but encounter issues that lead to diminished accuracy when applied to samples that disperse considerably. This work introduces an approach to split problematic sets of points using weighted k-means clustering. The multi-fidelity algorithm converges quickly on each of the smaller point clusters, leading to improved performance.

Chapter 2: Admissible Regions

2.1 Admissible Region Theory

Admissible regions were first introduced for the asteroid orbit determination problem, due to a similar challenge with short observation arcs (Milani et al., 2004). Researchers later extended the initial work into orbit determination for objects in Earth orbit (Tommei et al., 2007; Farnocchia et al., 2010). Broadly speaking, this problem can be characterized using a general nonlinear measurement model

$$\boldsymbol{z} = \boldsymbol{h}\left(\boldsymbol{x}; \boldsymbol{k}, t\right), \tag{2.1}$$

where $\boldsymbol{z} \in \mathbb{R}^m$ is the measurement vector, $\boldsymbol{x} \in \mathbb{R}^n$ is the state, $\boldsymbol{k} \in \mathbb{R}^l$ is a vector of parameters, and $t \in \mathbb{R}$ is the time (Worthy and Holzinger, 2015). For this problem, there is only one measurement or observation, meaning that $m < n$. This results in an under-determined system. The n states can be partitioned into observable and unobservable states

$$\boldsymbol{x}^T = \begin{bmatrix} \boldsymbol{x}_d^T & \boldsymbol{x}_u^T \end{bmatrix}, \tag{2.2}$$

where $\boldsymbol{x}_d \in \mathbb{R}^m$ are the determined states and $\boldsymbol{x}_u \in \mathbb{R}^u$ are the undetermined ones. Due to the inherent issue of observability, $\boldsymbol{x}_u$ has no impact on the actual measurement vector $\boldsymbol{y}$. As such, Eq. 2.1 can be written as

$$\boldsymbol{z} = \boldsymbol{h}\left(\boldsymbol{x}_d; \boldsymbol{k}, t\right). \tag{2.3}$$

Given a parameter vector $\boldsymbol{k}$ and time t, there exists a one-to-one and onto mapping from $\boldsymbol{x}_d$ to $\boldsymbol{y}$, meaning that Eq. 2.3 can be inverted into

$$\boldsymbol{x}_d = \boldsymbol{h}^{-1}\left(\boldsymbol{z}; \boldsymbol{k}, t\right). \tag{2.4}$$

To bound the undetermined state space $\mathbb{R}^u$, admissible region methods implement a set of constraint hypotheses $\boldsymbol{\mathcal{H}} = [\mathcal{H}_1 \ ... \ \mathcal{H}_c]$. These constraints can be defined using

a variety of methods, such as the dynamics of the underlying system or other physical limitations. They can be represented as

$$g_i\left(\boldsymbol{x}_d, \boldsymbol{x}_u; \boldsymbol{k}, t\right) \leq 0. \tag{2.5}$$

Substituting in Eq. 2.4 leads to

$$g_i\left(\boldsymbol{h}^{-1}(\boldsymbol{z}; \boldsymbol{k}, t), \boldsymbol{x}_u; \boldsymbol{k}, t\right) \leq 0. \tag{2.6}$$

These constraints are meaningful, as when an observation is first made, there exists an infinite continuum of solutions for $\boldsymbol{x}_u$. Eq. 2.6 reduces the solution space to a compact set. Formally, the set for constraint hypothesis $\mathcal{H}_i$ can be written as

$$\mathcal{R}_i = \{\boldsymbol{x}_u \in \mathbb{R}^u | \, g_i\left(\boldsymbol{h}^{-1}\left(\boldsymbol{z}; \boldsymbol{k}, t\right), \boldsymbol{x}_u; \boldsymbol{k}, t\right) \leq 0\}. \tag{2.7}$$

Each of these sets defines an admissible region in the undetermined state space. The boundary of an admissible region can be found by simply converting the inequality in Eq. 2.6 to an equal sign, leading to

$$g_i\left(\boldsymbol{h}^{-1}\left(\boldsymbol{y}; \boldsymbol{k}, t\right), \boldsymbol{x}_u; \boldsymbol{k}, t\right) = 0. \tag{2.8}$$

Any arbitrary number of admissible regions can be merged together. The combined region can be expressed as the intersection of the individual sets

$$\mathcal{R} = \mathcal{R}_1 \cap ... \cap \mathcal{R}_i \cap ... \cap \mathcal{R}_c, \tag{2.9}$$

where c denotes the number of constraints. For the problem of characterizing a space object's translation state, $n = 6$, as 6 pieces of independent information are needed. Traditionally, this can be in the form of Cartesian position and velocity or various orbital element sets. In the context of admissible regions, objects can be characterized in spherical coordinates relative to a sensor location. This spherical state can be written as

$$\boldsymbol{S} = \left[\rho, \dot{\rho}, \alpha, \delta, \dot{\alpha}, \dot{\delta}\right], \tag{2.10}$$

where ρ is the range from a sensor and $\dot{\rho}$ is the range-rate. α and δ are the angles of the object relative to the sensor-centered frame, and $\dot{\alpha}$ and $\dot{\delta}$ are the corresponding angle rates. Different definitions for the angles can be used, with common pairs including right ascension and declination, as well as azimuth and elevation. This work utilizes right ascension and declination, but the underlying theory applies regardless of definition with no loss of generality. As mentioned earlier, the challenge arises in that sensors in use today do not provide all elements of this state. Past research has focused on radar and optical sensors, which each provide different information. Radar sensors measure range, range-rate and angles, whereas optical sensors return values for the angles and angle rates. Therefore, the radar attributable is

$$\boldsymbol{z}_{rad} = [\rho, \dot{\rho}, \alpha, \delta]\,, \tag{2.11}$$

while the optical attributable is

$$\boldsymbol{z}_{opt} = \left[\alpha, \delta, \dot{\alpha}, \dot{\delta}\right]. \tag{2.12}$$

In each case, $m = 4$ and $u = 2$. There are two degrees of freedom, as there are two undetermined components of the state. These two unknowns form an infinite plane of solutions, which can be narrowed down using the previously discussed constraint hypotheses. One commonly applied constraint is based on the geocentric two-body energy per unit mass, which can be defined as

$$\varepsilon_E = \frac{1}{2}||\boldsymbol{v}||^2 - \frac{\mu_E}{||\boldsymbol{r}||}, \tag{2.13}$$

where ε_E is the specific energy, $\boldsymbol{r}$ and $\boldsymbol{v}$ are the object's geocentric position and velocity, and μ_E is the Earth's gravitational parameter. The object's position can be written as

$$\boldsymbol{r} = \boldsymbol{r}_o + \boldsymbol{r}_d, \tag{2.14}$$

where $\boldsymbol{r}_o$ is the position of the sensor and $\boldsymbol{r}_d = \rho\hat{\boldsymbol{r}}$ is the measurement vector pointing from the sensor to the object. Similarly, the velocity, $\boldsymbol{v}$, can be expressed as

$$\boldsymbol{v} = \boldsymbol{v}_o + \dot{\rho}\hat{\boldsymbol{r}} + \rho\dot{\alpha}\,\hat{\boldsymbol{r}}_\alpha + \rho\dot{\delta}\,\hat{\boldsymbol{r}}_\delta, \tag{2.15}$$

where $\boldsymbol{v}_o$ is the sensor's velocity, and $\hat{\boldsymbol{r}}_\alpha$ and $\hat{\boldsymbol{r}}_\delta$ are given by

$$\hat{\boldsymbol{r}}_{\boldsymbol{\alpha}} = \frac{\partial \hat{\boldsymbol{r}}}{\partial \alpha}, \qquad \hat{\boldsymbol{r}}_{\boldsymbol{\delta}} = \frac{\partial \hat{\boldsymbol{r}}}{\partial \delta}. \tag{2.16}$$

The three vectors $\hat{\boldsymbol{r}}, \hat{\boldsymbol{r}}_\alpha$, and $\hat{\boldsymbol{r}}_\delta$ define an orthogonal basis in 3-D space, and can be used to calculate the position and velocity in spherical coordinates. The magnitude of the position vector can be found using

$$||\boldsymbol{r}|| = \sqrt{\rho^2 + 2\,\rho\,\langle \boldsymbol{r}_o, \hat{\boldsymbol{r}} \rangle + ||\boldsymbol{r}_o||^2}. \tag{2.17}$$

Similarly, the magnitude of the object's velocity vector squared can be expressed as

$$\begin{aligned} ||\boldsymbol{v}||^2 = \dot{\rho}^2 + 2\,\dot{\rho}\,\langle \boldsymbol{v}_o, \hat{\boldsymbol{r}} \rangle + \rho^2 \left(\dot{\alpha}^2 \cos^2 \delta + \dot{\delta}^2 \right) \\ + 2\,\rho \left(\dot{\alpha} \langle \boldsymbol{v}_o, \hat{\boldsymbol{r}}_\alpha \rangle + \dot{\delta} \langle \boldsymbol{v}_o, \hat{\boldsymbol{r}}_\delta \rangle \right) + ||\boldsymbol{v}_o||^2, \end{aligned} \tag{2.18}$$

where $\langle \cdot \rangle$ denotes a dot product of two vectors. Eqs. 2.17 and 2.18 can be substituted into Eq. 2.13, resulting in an expression for the geocentric energy. For the optical case, the constraint boundary can be written as a function of the unknown variables, range and range-rate, leading to

$$2\,\varepsilon_E\,(\rho, \dot{\rho}) - \dot{\rho}^2 - 2\,\dot{\rho}\,\langle \boldsymbol{v}_o, \hat{\boldsymbol{r}} \rangle - T(\rho) \frac{2\,\mu_E}{\sqrt{S(\rho)}} = 0, \tag{2.19}$$

where

$$S(\rho) = \rho^2 + 2\,\rho\,\langle \boldsymbol{r}_o, \hat{\boldsymbol{r}} \rangle + ||\boldsymbol{r}_o||^2 \tag{2.20}$$

and

$$T(\rho) = \rho^2 \left(\dot{\alpha}^2 \cos^2 \delta + \dot{\delta}^2 \right) + 2\,\rho \left(\dot{\alpha} \langle \boldsymbol{v}_o, \hat{\boldsymbol{r}}_\alpha \rangle + \dot{\delta} \langle \boldsymbol{v}_o, \hat{\boldsymbol{r}}_\delta \rangle \right) + ||\boldsymbol{v}_o||^2. \tag{2.21}$$

Eq. 2.19 describes the relationship between an object's specific energy and various range and range-rate values. An orbit's specific energy is directly related to its semi-major axis a as

$$\varepsilon_E = -\frac{\mu_E}{2\,a}. \tag{2.22}$$

This indicates that orbital energy and semi-major axis are inversely related for the constraint equations. An example constraint hypothesis is that the object of interest

is under Earth's gravitational influence. For this to hold, the greatest specific energy the object could have would be $\varepsilon_E = 0$, which would correspond to a parabolic orbit. Another hypothesis could be a lower bound on the semi-major axis. A minimum energy, ε_{min}, corresponding to the minimum semi-major axis can be calculated using Eq. 2.22, and then applied to Eq. 2.19 to restrict the range and range-rate. Upon determining a value for ε_E, Eq. 2.19 can be converted into an inequality, with the sign indicating if the value acts as a lower or upper bound for the admissible region.

For the radar case, the corresponding attributable, $\boldsymbol{y}_{rad}$, contains information on range, range-rate, and angles, meaning that the angle rates are the unknowns. Eq. 2.19 can be expanded and rearranged into

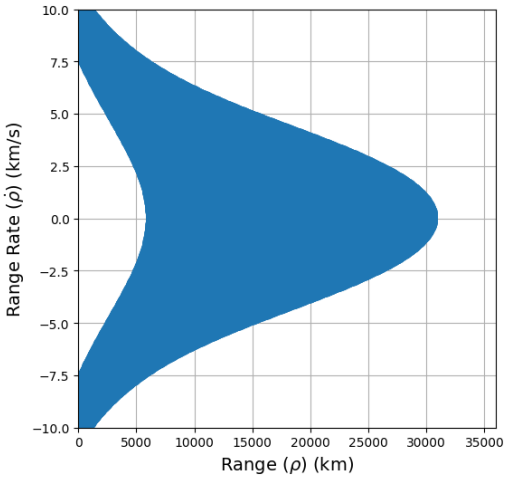
$$2\varepsilon_E\left(\dot{\alpha},\dot{\delta}\right) - \rho^2\cos^2\delta\,\dot{\alpha}^2 - \rho^2\,\dot{\delta}^2 - 2\rho\,\langle \boldsymbol{v}_o, \hat{\boldsymbol{r}}_\alpha\rangle\,\dot{\alpha} - 2\rho\,\langle \boldsymbol{v}_o, \hat{\boldsymbol{r}}_\delta\rangle\,\dot{\delta} - Z(\rho,\dot{\rho}) = 0, \quad (2.23)$$

where

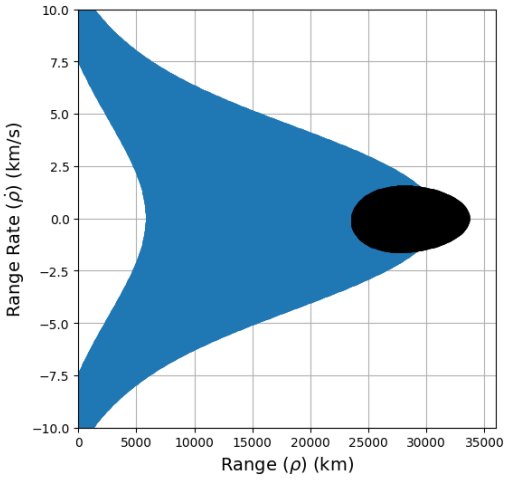
$$Z(\rho,\dot{\rho}) = \dot{\rho}^2 + 2\,\dot{\rho}\,\langle \boldsymbol{v}_o, \hat{\boldsymbol{r}}\rangle + ||\boldsymbol{v}_o||^2 - \frac{2\mu_E}{\sqrt{S(\rho)}}. \quad (2.24)$$

$S\,(\rho)$ is defined in Eq. 2.20. Just like in the optical case, a similar equation can be applied to the specific orbital energy, restricting the infinite set of solutions in the $\dot{\alpha}$-$\dot{\delta}$ plane to a more tractable region. The definitions of these constraints remain flexible as well, and are highly problem-dependent. For instance, operators interested in an object that is potentially reentering can select to not apply a minimum semi-major axis constraint. Operators investigating an object presumed to be in geostationary orbit can define tighter constraints for that orbit regime as well.

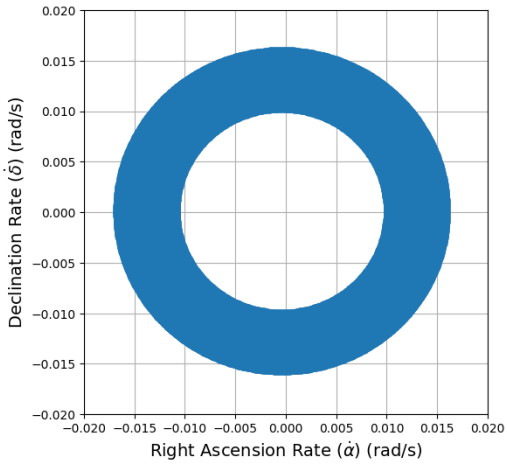
Figure 2.1 shows example admissible regions with different constraint hypotheses for both optical and radar sensors. Figure 2.1(a) presents semi-major axis (a) constraints for an optical case. The boundaries appear quadratic in the $\rho - \dot{\rho}$ plane. Figure 2.1(b) illustrates an additional eccentricity (e) constraint. Figure 2.1(c) depicts semi-major axis constraints for the radar case, with circular boundaries in the $\dot{\alpha} - \dot{\delta}$ space. Figure 2.1(d) introduces radius of periapsis (r_p) constraints, which are hyperbolic in this case. These constraint hypotheses are by no means comprehensive.



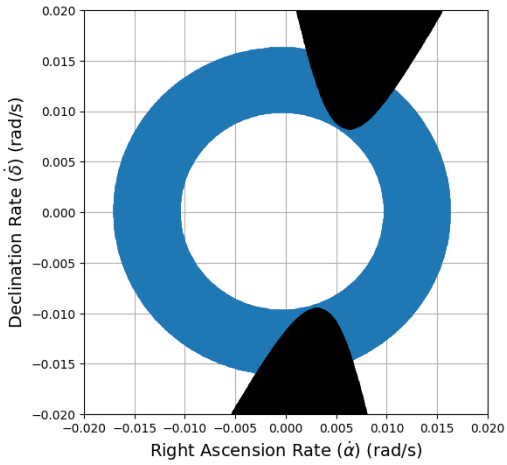
(a) Optical with a constraint (blue)



(b) Optical with a constraint (blue) and e constraint (black)



(c) Radar with a constraint (blue)



(d) Radar with a constraint (blue) and r_p constraint (black)

Figure 2.1: Example Admissible Regions

Additional constraint hypotheses can be added and combined into a total admissible region. The selection and definition of constraints is dependent on the specific problem and the assumptions one makes about a newly detected object. Once an admissible region is generated, it needs to be propagated forward in time, as the space object has continued moving after the initial observation. As such, this research leverages

some advancements in uncertainty propagation to do so quickly and effectively.

2.2 Admissible Region Representation

Since some admissible regions are analytically defined by continuous functions, an issue arises in discretizing them for practical applications. Initially, possible approaches to sampling the continuous admissible regions included Delaunay triangulation for the optical case, and a cobweb approach for the radar case (Tommei et al., 2007). Researchers also transformed admissible regions into Delaunay variables and used intersection theory to associate separate observations (Maruskin et al., 2009).

More recently, researchers have considered representing the admissible region probabilistically. This makes sense, as the true state of a space object remains unknown within the admissible region space. Some have proposed mapping the admissible region into 6-D Poincaré space and then defining probability density functions (PDFs) over individual bins (Fujimoto and Scheeres, 2012). Researchers have acknowledged the computational challenges of this type of approach, and have tried to address them using linear approximations during the mapping process. In a separate work, Bayesian admissible regions were utilized for data association in low Earth orbit (LEO) (Fujimoto and Scheeres, 2013).

Additional work has been done for stochastically approximating an admissible region using a Gaussian mixture (DeMars and Jah, 2013) The advantages of doing so were demonstrated by forecasting and updating the approximated PDF with subsequent data to iteratively refine the admissible region. Others have also analyzed the relationships between different types of uncertainty and their impacts on a probabilistic admissible region (Worthy and Holzinger, 2015). These advancements were then leveraged as a way to initialize orbits for multi-hypothesis filtering (Hussein et al., 2018; Schumacher et al., 2018).

More recently, the Gaussian mixture approach has been further refined to represent probabilistic admissible regions (Zucchelli et al., 2021). This work utilizes

a Gaussian mixture method to instantiate PDFs for admissible regions. The mean values for the Gaussian components are defined based on a grid in the two unknown dimensions (i.e., range and range-rate for an optical case, angle rates for a radar case). The covariances in the unknown dimensions are defined based on the grid spacing such that the overall Gaussian mixture is near-uniform over the entire admissible region. This is because without additional *a priori* information, there is an equal probability of the space object's true state falling anywhere within the admissible region. This process can generate a 2-D Gaussian mixture model that represents the PDF of any given admissible region.

Chapter 3: Multi-Fidelity Propagation

In order to model how a generated Gaussian mixture evolves over time, an approach for uncertainty propagation is required. For this work, each individual Gaussian component, with a mean and covariance, is converted into a set of sigma points via an unscented transform. These individual sigma points are then propagated forward in time using nonlinear orbital dynamics and recombined for an estimate of the admissible region PDF.

The next area of focus then becomes propagation fidelity. A low-fidelity propagator with only two-body dynamics and J_2 is fairly computationally cheap to run, but results in larger errors. On the other hand, a high-fidelity propagator that takes atmospheric drag, solar radiation pressure, third-body effects, etc. into account has a more accurate output, but is much more costly to calculate. In general, there is a direct tradeoff between propagation accuracy and computational cost.

Researchers in the field of scientific computing have found a way to bypass this tradeoff and developed various multi-fidelity propagation algorithms that perform better than a low-fidelity propagator without the cost of a high-fidelity one (Narayan et al., 2014). This work leverages one such algorithm, known as stochastic collocation, as it has already been adopted for orbital dynamics models in previous works (Jones and Weisman, 2019; Zucchelli et al., 2021). Consider a set of points $\boldsymbol{\xi} \in \mathbb{R}^{n \times p}$ that need to be propagated forward for $\boldsymbol{t} \in \mathbb{R}^{q}$ times. This propagation can be performed with a low-fidelity propagator, resulting in

$$\boldsymbol{Y}^{L} = \left[\boldsymbol{y}^{L}\left(\boldsymbol{\xi}_{1}\right) \ ... \ \boldsymbol{y}^{L}\left(\boldsymbol{\xi}_{p}\right)\right], \tag{3.1}$$

where $\boldsymbol{Y}^{L} \in \mathbb{R}^{(nq) \times p}$ contains the states of the points at every time step of interest. Moving forward, the L superscript indicates samples processed with a low-fidelity propagator. Next, important samples are selected from $\boldsymbol{Y}^{L}$ using a pivoted Cholesky

decomposition

$$\left(\boldsymbol{Y}^L\right)^T \boldsymbol{G}^L \left(\boldsymbol{Y}^L\right) = \boldsymbol{A}^T \boldsymbol{L}\boldsymbol{L}^T \boldsymbol{A}, \tag{3.2}$$

where $\boldsymbol{G}^L$ is the Gramian matrix of the low-fidelity points and $\boldsymbol{A}$ is a pivot matrix that orders samples according to their distance from the space formed by $\boldsymbol{Y}^L$. A full description of how to calculate the $\boldsymbol{L}$ matrix and the important samples can be found in Narayan et al. (2014). The algorithm then attempts to approximate each low-fidelity sample $\boldsymbol{y}^L$ with a product of a set of coefficients and the important samples, shown as

$$\boldsymbol{y}^L(\boldsymbol{\xi}) \approx \hat{\boldsymbol{y}}^L(\boldsymbol{\xi}) = \sum_{i=1}^{r} c_i(\boldsymbol{\xi})\, \boldsymbol{y}^L(\bar{\boldsymbol{\xi}}_i), \tag{3.3}$$

where $\hat{\boldsymbol{y}}^L$ is an approximated sample, $\boldsymbol{c} = [c_1\left(\boldsymbol{\xi}\right) \ ... \ c_r\left(\boldsymbol{\xi}\right)]$ are the coefficients, and $\boldsymbol{y}^L(\bar{\boldsymbol{\xi}}_i)$ are the important samples. The number of important samples, r, is iteratively increased until the approximation error meets some defined tolerance. This is expressed as

$$||\boldsymbol{Y}^L - \hat{\boldsymbol{Y}}^L|| \leq \boldsymbol{\varepsilon}_{tol}. \tag{3.4}$$

The initial conditions for the important samples, $\bar{\boldsymbol{\xi}}$, are then re-propagated using a high-fidelity model, resulting in $\boldsymbol{Y}^H\left(\bar{\boldsymbol{\xi}}\right)$. The H superscript refers to high-fidelity points. These high-fidelity points are substituted back into Eq. 3.3 to find multi-fidelity estimates for each of the samples, written as

$$\hat{\boldsymbol{y}}^H(\boldsymbol{\xi}) = \sum_{i=1}^{r} c_i(\boldsymbol{\xi})\, \boldsymbol{y}^H(\bar{\boldsymbol{\xi}}_i). \tag{3.5}$$

This process creates multi-fidelity samples $\hat{\boldsymbol{Y}}^H = \left[\hat{\boldsymbol{y}}^H\left(\boldsymbol{\xi}_1\right) \ ... \ \hat{\boldsymbol{y}}^H\left(\boldsymbol{\xi}_p\right)\right]$ with a limited number of high-fidelity propagations. The benefits of this approach become apparent with more points. An admissible region Gaussian mixture can have thousands of sigma points. Running a full, high-fidelity propagator for these points is very time-consuming, whereas a low-fidelity one would result in a significant amount of error. This multi-fidelity approach only performs high-fidelity propagation for a small set of samples, but is able to leverage that knowledge to correct all of the low-fidelity points and improve the model's overall accuracy.

Chapter 4: Admissible Region Splitting

4.1 Problem Statement

A challenge arises in that admissible regions with more modest constraints spread out quite rapidly as they are propagated over time. This rapid growth creates issues with the multi-fidelity propagator, as the correction calculated from the high-fidelity points is inadequate for correcting the highly disparate set of low-fidelity points. Recall from Eqs. 3.3 and 3.4 that the stochastic collocation algorithm seeks to find an approximation within some error threshold ε_{tol} with r important samples. Large initial regions can fail to reach this threshold with fewer than r_{max} important samples.

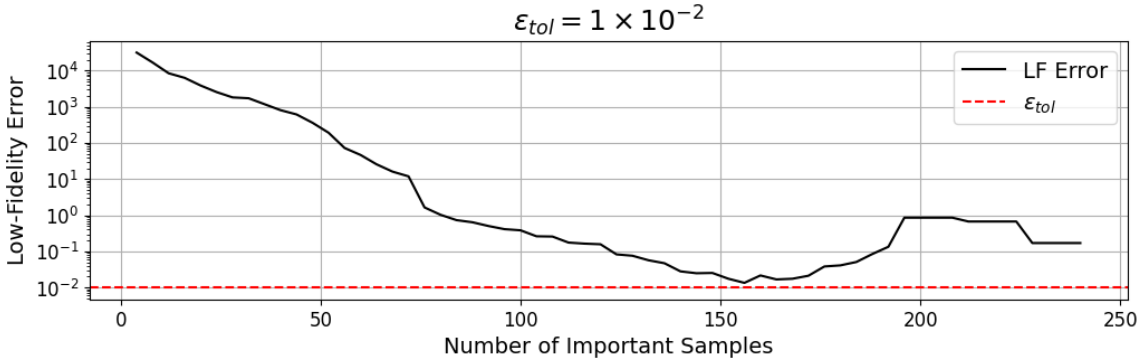


Figure 4.1: Example of Low-Fidelity Overfitting

Figure 4.1 illustrates one such instance. It shows that the low-fidelity error can fail to reach the error threshold, and can actually increase as the approximation becomes overfitted. In this case, the individual points are too far dispersed, and cannot be accurately approximated with a limited number of important samples. To deal with this issue, this work proposes a novel method for splitting an admissible region into smaller regions. The multi-fidelity propagation can then be performed on each sub-problem.

4.2 Splitting Methodology

Past works have considered using semi-major axis as a splitting parameter for an admissible region (Zucchelli et al., 2021). This is because semi-major axis is intuitively the most important factor when considering sample dispersion. However, this approach ignores the information contained within the other orbital elements. A naive approach would then be to split along each orbital element dimension. This in turn neglects the fact that each orbital element has different contributions to sample dispersion. Empirical tests show that splitting using all orbital elements has minimal or even detrimental effects on the multi-fidelity propagator. To further investigate the impact of each orbital element, their partial derivatives with respect to the admissible region parameters are taken. The Keplerian orbital elements can be defined as

$$\boldsymbol{K} = [a, e, i, \Omega, \omega, \theta], \tag{4.1}$$

where a is the semi-major axis, e is the eccentricity, i is the inclination, Ω is the right angle of the ascending node, ω is the argument of periapsis, and θ is the true anomaly. These Keplerian elements can be expressed as functions of Cartesian position and velocity (Vallado, 2022). As such, partial derivatives can be taken with respect to the Cartesian state vector $\boldsymbol{C}$, written as

$$\frac{\partial \boldsymbol{K}}{\partial \boldsymbol{C}} = \begin{bmatrix} \dfrac{\partial a}{\partial r_x} & \cdots & \dfrac{\partial a}{\partial v_z} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial \theta}{\partial r_x} & \cdots & \dfrac{\partial \theta}{\partial v_z} \end{bmatrix}, \tag{4.2}$$

where

$$\boldsymbol{C} = [r_x, r_y, r_z, v_x, v_y, v_z]. \tag{4.3}$$

Furthermore, $\boldsymbol{C}$ can be analytically expressed in spherical coordinates. Therefore, partials of $\boldsymbol{C}$ with respect to the spherical state vector $\boldsymbol{S}$ defined in Eq. 2.10 can be

found as

$$\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{S}} = \begin{bmatrix} \frac{\partial r_x}{\partial \rho} & \cdots & \frac{\partial r_x}{\partial \dot{\delta}} \\ \vdots & \ddots & \vdots \\ \frac{\partial v_z}{\partial \rho} & \cdots & \frac{\partial v_z}{\partial \dot{\delta}} \end{bmatrix}. \tag{4.4}$$

The partials of $\boldsymbol{K}$ with respect to $\boldsymbol{S}$ can be found as the matrix product of these two matrices. This results in

$$\frac{\partial \boldsymbol{K}}{\partial \boldsymbol{S}} = \frac{\partial \boldsymbol{K}}{\partial \boldsymbol{C}} \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{S}} = \begin{bmatrix} \frac{\partial a}{\partial r_x}\frac{\partial r_x}{\partial \rho} + \ldots + \frac{\partial a}{\partial v_z}\frac{\partial v_z}{\partial \rho} & \cdots & \frac{\partial a}{\partial r_x}\frac{\partial r_x}{\partial \dot{\delta}} + \ldots + \frac{\partial a}{\partial v_z}\frac{\partial v_z}{\partial \dot{\delta}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \theta}{\partial r_x}\frac{\partial r_x}{\partial \rho} + \ldots + \frac{\partial \theta}{\partial v_z}\frac{\partial v_z}{\partial \rho} & \cdots & \frac{\partial \theta}{\partial r_x}\frac{\partial r_x}{\partial \dot{\delta}} + \ldots + \frac{\partial \theta}{\partial v_z}\frac{\partial v_z}{\partial \dot{\delta}} \end{bmatrix}. \tag{4.5}$$

In essence, this matrix captures the relationship between changes in a sample's spherical state and its Keplerian elements. Given an optical measurement, the values for $\alpha, \delta, \dot{\alpha}$ and $\dot{\delta}$ are available, with some variance due to sensor noise and errors. However, the admissible region approach creates samples that span a wide array of ρ and $\dot{\rho}$ values. The partials with respect to ρ and $\dot{\rho}$ can be numerically evaluated using samples drawn from the admissible region. The values for semi-major and eccentricity cannot be easily compared, due to the mismatched units, but empirical tests reveal that they still contribute to sample dispersion as well. Semi-major axis helps determine the size of a point's orbit, whereas the eccentricity characterizes the orbit's shape. The angles, on the other hand, can be compared, and the results for an optical case are shown in Table 4.1.

Table 4.1: Average Optical Partials

Orbital Element	i	Ω	ω	θ
$\frac{\partial}{\partial \rho}\left(\frac{\text{rad}}{\text{km}}\right)$	1.3×10^{-6}	3.2×10^{-6}	1.5×10^{-4}	1.5×10^{-4}
$\frac{\partial}{\partial \dot{\rho}}\left(\frac{\text{rad} \cdot \text{s}}{\text{km}}\right)$	1.5×10^{-3}	2.8×10^{-3}	0.459	0.987

It becomes apparent that θ and ω are the most affected, followed by Ω and i. This makes sense, as knowing the angles and angle rates provides excellent information about the location and orientation of the object's orbit plane. This is reflected in the magnitudes of the partials for inclination and right ascension of the ascending node. However, there is ambiguity in terms of the size, shape, and orientation of the object's orbit within that plane. A similar analysis can be performed for the radar case, but relative to the unknowns $\dot{\alpha}$ and $\dot{\delta}$. The results are shown in Table 4.2.

Table 4.2: Average Radar Partials

Orbital Element	i	Ω	ω	θ
$\frac{\partial}{\partial \dot{\alpha}}$ (s)	5.5×10^1	3.8×10^1	2.5×10^3	2.5×10^3
$\frac{\partial}{\partial \dot{\delta}}$ (s)	1.1×10^2	2.3×10^2	1.2×10^3	1.1×10^3

A similar pattern can be seen with θ and ω being the largest in magnitude. A radar observation provides plentiful information about an object's position in space, but much less regarding the direction it is traveling in. The wide array of values for the angle rates have an impact across all of the samples' orbital elements.

The splitting algorithm based solely on semi-major axis is effective, as it is still a major factor that drives sample dispersion, but incomplete. In optical cases, it can group together points that spread out along the orbit plane, whereas in the radar case, it can create regions that are traveling in completely different directions. A weighted k-means clustering algorithm based on a, ω, and θ is utilized to segment the samples into individual regions. K-means follows an expectation maximization approach to clustering and minimizes the distance of each sample from its corresponding cluster centroid (MacQueen, 1967). The features are weighted according to the numerically calculated partials. The exact values of the weights can vary, but in general, the semi-major axis values are weighted an order of magnitude greater than the angular values. An example set of weights could be $w_a = 10$, $w_\omega = 1.2$ and $w_\theta = 1.1$.

An issue arises in that k-means uses a Euclidean distance-based metric, meaning that angular values, like ω or θ, can introduce problems. For instance, angular values of 179° and -179° are only 2° apart, but on a linear scale, the distance between them is 358°. To deal with this, the angular values are projected onto the unit circle by taking the sine and cosine of each value. This helps deal with these angular uncertainties, as similar angles are located close to each other along a 2-D unit circle. The weights are adjusted accordingly as well.

Performing this clustering over multiple parameters leads to improvements in runtime and the number of regions needed to achieve an accuracy threshold. This is because the multi-fidelity approach performs better for samples that remain closer together in phase space. Grouping points by semi-major axis captures some of this spatial behavior, but including additional information on each sample's Keplerian elements leads to further improvements. In turn, this benefits rapid sensor tasking when considering large initial admissible regions.

4.3 Numerical Tests

This section presents different test cases of newly detected objects used to verify the performance of the clustering methodology discussed earlier.

4.3.1 Optical Measurement

This section presents an optical test case. The measurement attributable is

$$\boldsymbol{z}_{opt} = \left[-0.597097,\ -0.090847,\ 9.637586 \times 10^{-5},\ 2.467493 \times 10^{-5}\right]. \tag{4.6}$$

The angles have units of radians, and the angle rates have units of radians per second. The simulated station position and velocity are

$$\boldsymbol{r}_o = [5588.520,\ -3018.76,\ -579.802]\ \text{km} \tag{4.7}$$

and

$$\boldsymbol{v}_o = [0.220132,\ 0.407521,\ 0.000]\ \text{km/s}. \tag{4.8}$$

Constraints are then applied to this admissible region. The semi-major axis has a minimum value of $R_E + 2500$ km and a maximum value of 40000 km, where R_E is the radius of the Earth. The radius of periapsis has a minimum value of $R_E + 2500$ km as well.

The grid is then sampled in range increments of 100 km and range-rate increments of 0.1 km/s. This results in a total of 3454 points that meet the constraints. These points are used as the mean vector for the Gaussian components. The variances for the range (σ_ρ) and range-rate ($\sigma_{\dot{\rho}}$) are defined as 3 times the grid spacing, meaning that $\sigma_\rho = 300$ km and $\sigma_{\dot{\rho}} = 0.3$ km/s, respectively. The measurement variances are 1 arcsecond for the angles (α, δ), and 0.1 arcseconds/s for the angle rates $\left(\dot{\alpha}, \dot{\delta}\right)$. The unscented transform creates $2n + 1$ sigma points for each Gaussian component, leading to 44902 sigma points. Figure 4.2 shows the initial admissible region. The red

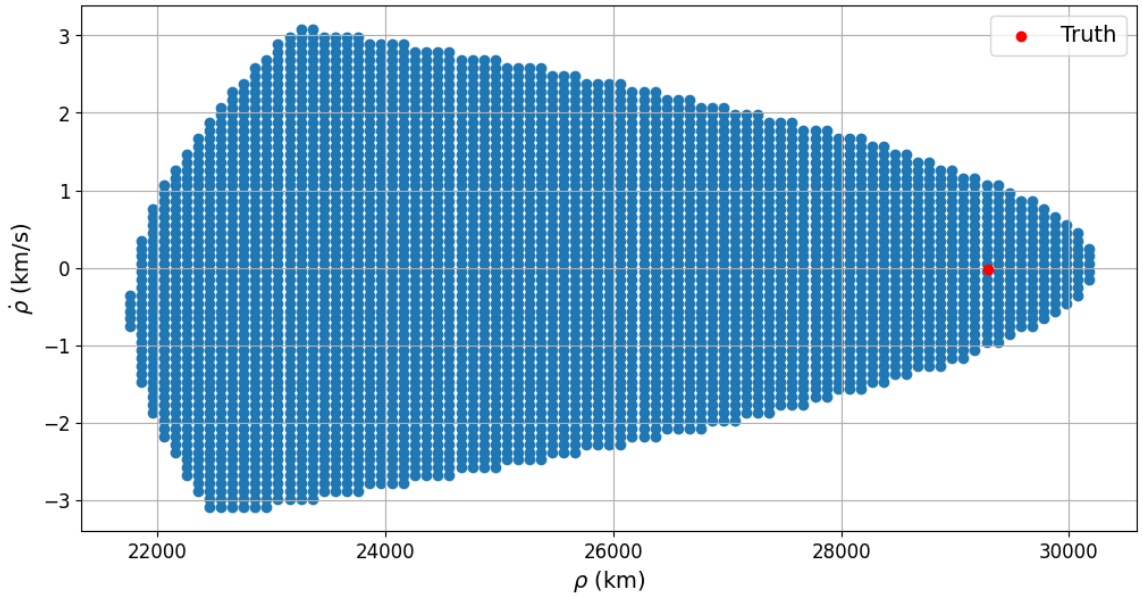


Figure 4.2: Initial Admissible Region in $\rho - \dot{\rho}$ Space

dot indicates the true value of the object. Each of the sigma points is then converted into Cartesian coordinates and propagated forward in time for 1 day in 30 minute intervals. The spacecraft parameters used for the high-fidelity force models are shown

in Table 4.3. The force models used for the low- and high-fidelity propagators are shown in Table 4.4.

Table 4.3: Spacecraft Parameters

Satellite Mass	Drag and SRP Area	C_D	C_R
500 kg	1 m^2	2.0	1.5

Table 4.4: Propagator Force Models

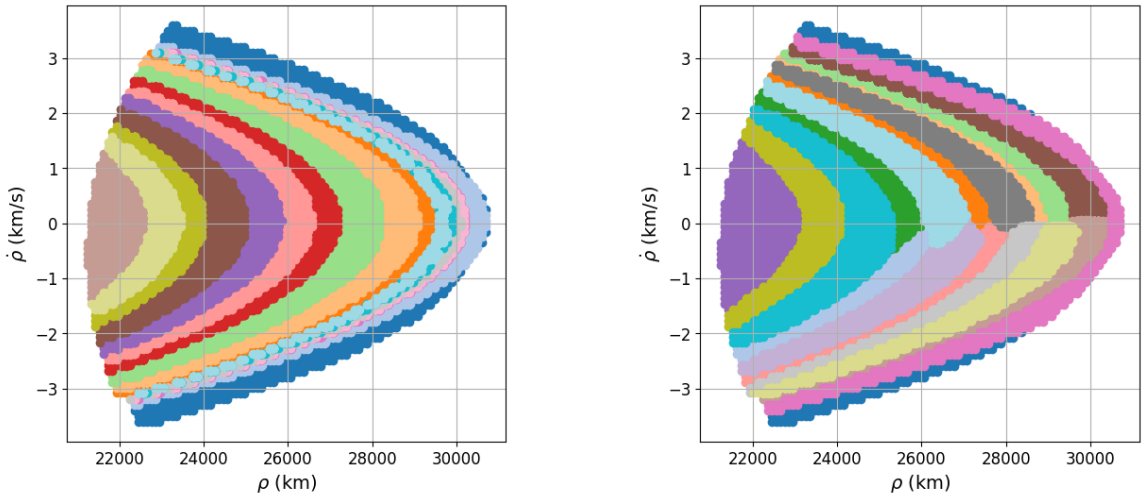
Force Model	Low-Fidelity	High-Fidelity
Central Body Gravity	Two-Body and J_2	70×70
Third-Body Perturbations	None	Sun, Moon
Solar Radiation Pressure (SRP)	None	Cannonball
Atmospheric Drag	None	Cannonball

It is also important to consider the runtime of these algorithms. The admissible region driver is written in Python, while the propagation code is in C. The propagations are performed serially for these tests, but this process can be parallelized to some degree to achieve performance improvements. The tests are performed on an AMD Ryzen 5 3600 CPU with 16 GB of RAM. Table 4.5 shows the final position root mean square error (RMSE) and runtime for three separate propagators, a high-fidelity one, a low-fidelity one, and a multi-fidelity one applied to all the points with a maximum of 100 important samples.

Table 4.5: Propagator Accuracy and Runtime Performance

Propagator	Final Position RMSE (km)	Runtime (min)
High-Fidelity	0.000	32.2
Low-Fidelity	6.449	6.0
Multi-Fidelity	7.282	7.2

The computation time taken for each of the cases meets our expectations, with the high-fidelity propagator being the most expensive, the low-fidelity one taking the least amount of time, and the multi-fidelity approach falling somewhere in the middle. In terms of RMSE, the high-fidelity propagator acts as a baseline, resulting in an error of 0. The low-fidelity propagator has moderate error values, whereas the multi-fidelity approach is actually less accurate. This indicates that this test case encounters the previously identified issue for large initial admissible regions. The sigma points drift far apart in phase space, meaning that the stochastic collocation algorithm is unable to find a set of important samples that can accurately represent all of the points. The admissible region therefore has to be split into several smaller regions with points that remain closer together over the course of propagation. The multi-fidelity algorithm can then be run on each individual region.



(a) Semi-major axis splitting with 17 regions

(b) K-means clustering with 18 regions

Figure 4.3: Split Admissible Regions

Figure 4.3 shows the regions formed by two splitting approaches, one that splits based on only semi-major axis, and a k-means clustering approach. The semi-major axis-based approach forms clearly identifiable, parabolic bands. The same pattern exists in the k-means clusters, but there is also separation along the $\dot{\rho} = 0$ line. Since

the primary cause of the multi-fidelity propagation failing is sample dispersion across the phase space, a good metric for the efficacy of splitting is the average distance between points within a region. Without any splitting, the average distance between points at the final time is 40665.5 km. The semi-major axis splitting method reduces the average distance to 27406.1 km within each region, and the clustering method is able to perform even better, with an average distance of 18493.5 km.

This improvement in terms of how close the sigma points remain is also reflected in the algorithm's accuracy. Figure 4.4 shows the RMSE of the multi-fidelity approach with splitting involved. The dashed red line indicates the accuracy of the low-fidelity samples, the dashed blue line denotes the accuracy of the semi-major axis splitting approach, and the black line illustrates the performance of the clustering method. As expected, when there are only a few regions, the earlier issue with sigma

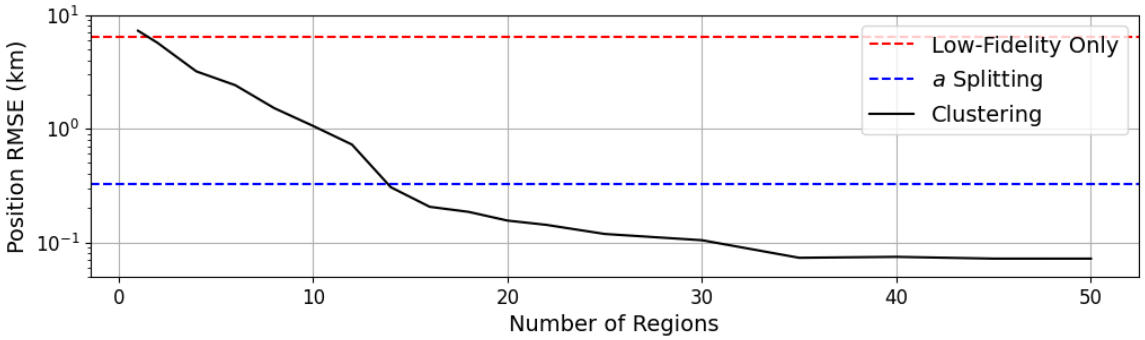


Figure 4.4: Effect of the Number of Regions on Clustering Performance

points being too sparse persists. However, with more and more clusters, the algorithm's performance improves and surpasses the a-only approach. However, it should be noted that clustering over more regions has diminishing returns. As the sum of ω and θ is the argument of latitude, clustering over the semi-major axis and argument of latitude is also performed. Doing so results in a slight decrease in performance, indicating that there is value in clustering with the two angles independently. The semi-major axis splitting approach resulted in final position errors of several hundred meters, whereas the clustering approach was able to reduce it to the order of tens

of meters. Figure 4.5 shows the runtime of these methods. The clustering approach

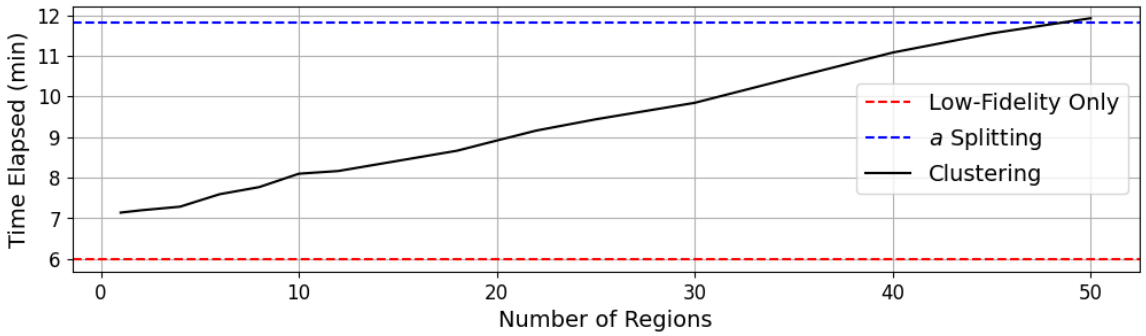


Figure 4.5: Effect of the Number of Regions on Runtime

requires less time with a limited number of clusters. It starts to take longer once too many clusters are included, but that value is reached far beyond the number necessary to surpass the a-only splitting approach in terms of accuracy. The semi-major axis splitting approach is recursive, which allows it to focus on groups of samples that are problematic for the multi-fidelity algorithm. However, this comes at a cost, as the important sample selection process has to be continuously called. On the other hand, k-means clustering is relatively efficient, and only needs to be called initially to define the individual regions.

4.3.2 Radar Measurement

This section presents a radar test case. The measurement attributable is

$$\boldsymbol{z}_{rad} = [1122.615,\ -2.396970,\ -0.380677,\ -0.131821]\,. \tag{4.9}$$

The range has units of km, the range-rate has units of km/s, and the angles have units of radians. The simulated station position and velocity are

$$\boldsymbol{r}_o = [6315.292,\ 562.0487,\ -694.125]\ \text{km} \tag{4.10}$$

and

$$\boldsymbol{v}_o = [-0.04099,\ 0.460518,\ 0.000]\ \text{km/s}. \tag{4.11}$$

Constraints are then applied to this admissible region. The semi-major axis has a minimum value of $R_E + 500$ km and a maximum value of $R_E + 2500$ km, where R_E is the radius of the Earth. The radius of periapsis has the same constraints as well.

The grid is then sampled in increments of 1×10^{-4} rad/s for both $\dot{\alpha}$ and $\dot{\delta}$. This results in a total of 813 points that meet the constraints. These points are used as the mean vector for the Gaussian components. The variances for the angle rates $(\sigma_{\dot{\alpha}}, \sigma_{\dot{\delta}})$ are defined as 3 times the grid spacing, meaning that $\sigma_{\dot{\alpha}} = \sigma_{\dot{\delta}} = 3 \times 10^{-4}$ rad/s. The measurement variances are 0.1 km for the range, 1 m/s for the range-rate, and 1 arcsecond for the angles. The unscented transform creates $2n + 1$ sigma points for each Gaussian component, leading to 10569 sigma points.

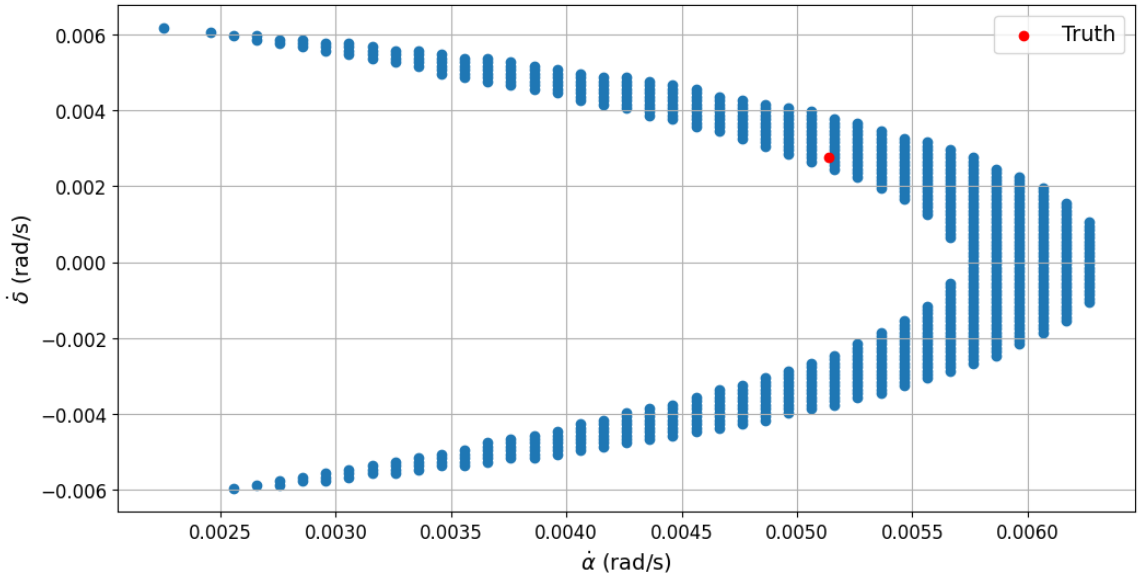


Figure 4.6: Initial Admissible Region in $\dot{\alpha} - \dot{\delta}$ Space

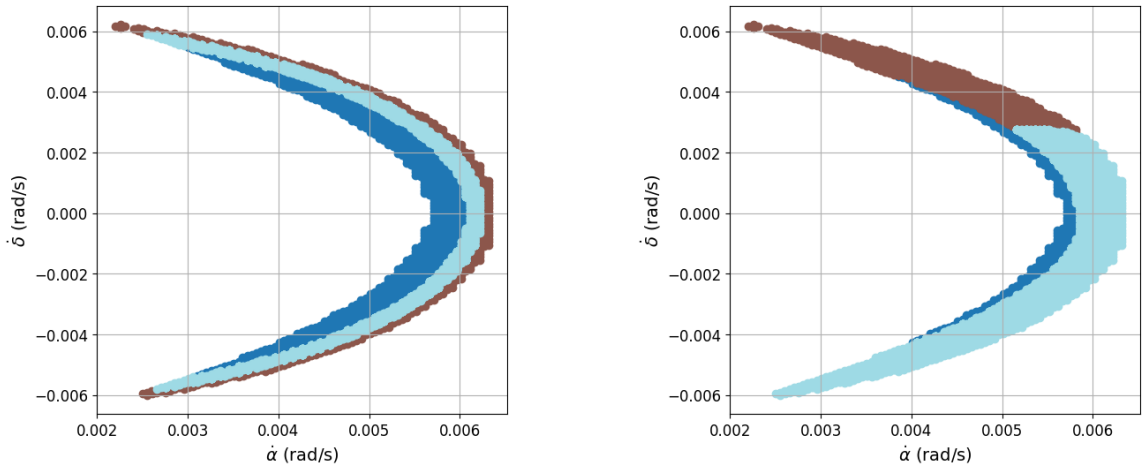
Figure 4.6 shows the initial admissible region. The red dot indicates the true value of the object. Each of the sigma points is then converted into Cartesian coordinates and propagated forward in time for 1 day in 30 minute intervals. The spacecraft parameters and force models are the same as the previous test. Table 4.6 shows the final position root mean square error (RMSE) and runtime for the three separate

propagators.

Table 4.6: Propagator Accuracy and Runtime Performance

Propagator	Final Position RMSE (km)	Runtime (min)
High-Fidelity	0.000	30.4
Low-Fidelity	1.083	4.9
Multi-Fidelity	0.497	6.3

In this case, the base multi-fidelity algorithm applied to the entire region provided some accuracy improvements with a small increase in runtime. The different splitting algorithms are then applied to this initial region.



(a) Semi-major axis splitting with 3 regions

(b) K-means clustering with 3 regions

Figure 4.7: Split Admissible Regions

Figure 4.7 shows the regions formed by the two splitting approaches. The average distances between points within each region are computed as well. Initially, the average distance between points is 10601.3 km. The semi-major axis splitting method reduces the average distance to 10165.4 km within each region, and the clustering method is able to perform slightly better, with an average distance of 9345.1 km.

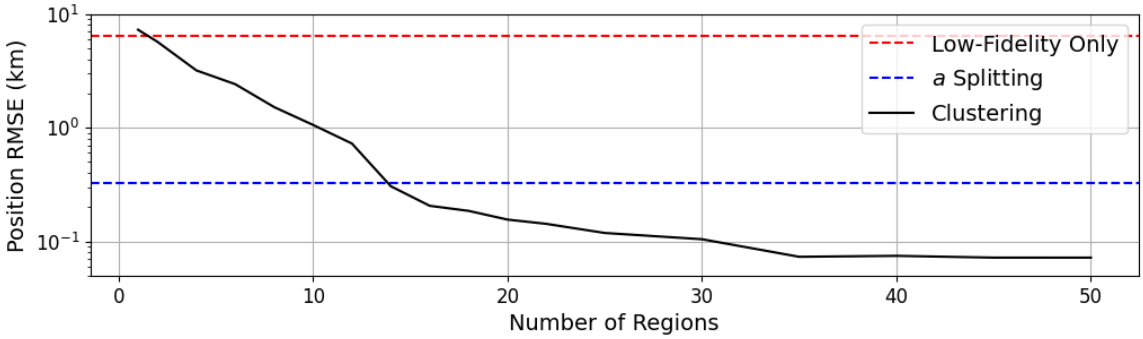


Figure 4.8: Effect of the Number of Regions on Clustering Performance

These improvements are also reflected in this test's accuracy. Figure 4.8 shows the RMSE of the multi-fidelity approach with splitting involved. The dashed red line indicates the accuracy of the low-fidelity samples, the dashed blue line denotes the accuracy of the semi-major axis splitting approach, and the black line illustrates the performance of the clustering method. The semi-major axis splitting approach results in final position errors of tens of meters, whereas the clustering approach is able to reduce it to the order of meters.

Runtime tests are also performed using the same software and hardware detailed in the previous section. A similar pattern is observed, with the k-means clus-

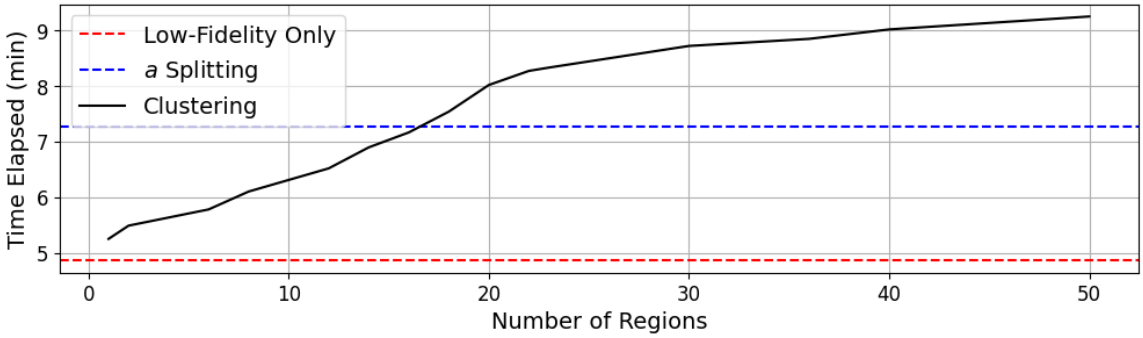


Figure 4.9: Effect of the Number of Regions on Runtime

tering approach requiring less time with fewer clusters. Both splitting approaches are more expensive than the low-fidelity model, but offer accuracy improvements.

Chapter 5: Conclusion

Admissible region approaches have risen in popularity as a way to actively perform initial orbit determination. This work utilized multi-fidelity propagation to quickly predict the evolution of an admissible region over time. For large initial admissible regions with loose constraints, however, the multi-fidelity algorithm can fail. To deal with this, this work developed a novel method for splitting an admissible region into smaller regions that leverages information about each point's orbital elements. Partial derivatives of the Keplerian orbital elements with respect to an object's spherical state were taken and used to select which orbital elements to utilize. The samples were clustered using a k-means clustering algorithm, and the impact of each parameter on the multi-fidelity model's performance was characterized. The method's benefits were validated using numerical examples for both optical and radar cases, which demonstrated that clustering using this additional information led to improvements in accuracy and runtime compared to splitting solely using semi-major axis.

An avenue for future work could be analyzing how to discretize the admissible region further. For this work, the grid used to generate the Gaussian components was uniformly spaced, but including knowledge about how the samples evolve could inform a more educated arrangement of components. Another possibility is analyzing different clustering algorithms. K-means is a centroid-based algorithm that takes the number of clusters as an input, whereas other clustering methods can determine that value independently.

Works Cited

K. J. DeMars and M. K. Jah. Probabilistic initial orbit determination using gaussian mixture models. *Journal of Guidance, Control, and Dynamics*, 36(5): 1324–1335, 2013. doi: https://doi.org/10.2514/1.59844.

D. Farnocchia, G. Tommei, A. Milani, and A. Rossi. Innovative methods of correlation and orbit determination for space debris. *Celestial Mechanics and Dynamical Astronomy*, 107:169–185, 2010. doi: https://doi.org/10.1007/s10569-010-9274-6.

K. Fujimoto and D. J. Scheeres. Correlation of optical observations of earth-orbiting objects and initial orbit determination. *Journal of Guidance, Control, and Dynamics*, 35(1):208–221, 2012. doi: https://doi.org/10.2514/1.53126.

K. Fujimoto and D. J. Scheeres. Applications of the admissible region to space-based observations. *Advances in Space Research*, 52(4):696–704, 2013. doi: https://doi.org/10.1016/j.asr.2013.04.020.

I. I. Hussein, C. W. T. Roscoe, M. Mercurio, M. P. Wilkins, and P. W. Schumacher. Probabilistic admissible region for multihypothesis filter initialization. *Journal of Guidance, Control, and Dynamics*, 41(3):710–724, 2018. doi: https://doi.org/10.2514/1.G002788.

B. A. Jones and R. Weisman. Multi-fidelity orbit uncertainty propagation. *Acta Astronautica*, 155:406–417, 2019. doi: https://doi.org/10.1016/j.actaastro.2018.10.023.

A. Kalur, S. A. Szklany, and J. L. Crassidis. Space object data association using spatial pattern recognition approaches. *The Journal of the Astronautical Sciences*, 67:1708–1734, 2020. doi: https://doi.org/10.1007/s40295-020-00217-0.

J. B. MacQueen. Some methods for classification and analysis of multivariate observations, 1967.

J. M. Maruskin, D. J. Scheeres, and K. T. Alfriend. Correlation of optical observations of objects in earth orbit. *Journal of Guidance, Control, and Dynamics*, 32(1):194–209, 2009. doi: https://doi.org/10.2514/1.36398.

A. Milani, G. F. Gronchi, M. D. Vitturi, and Z. Knežević. Orbit determination with very short arcs. i admissible regions. *Celestial Mechanics and Dynamical Astronomy*, 90:57–85, 2004. doi: https://doi.org/10.1007/s10569-004-6593-5.

A. Narayan, C. Gittelson, and D. Xiu. A stochastic collocation algorithm with multifidelity models. *SIAM Journal on Scientific Computing*, 36(2):495–521, 2014. doi: https://doi.org/10.1137/130929461.

P. W. Schumacher, J. A. Gaebler, C. W. T. Roscoe, M. P. Wilkins, and P. Axelrad. Parallel initial orbit determination using angles-only observation pairs. *Celestial Mechanics and Dynamical Astronomy*, 130(60), 2018. doi: https://doi.org/10.1007/s10569-018-9852-6.

G. Tommei, A. Milani, and A. Rossi. Orbit determination of space debris: Admissible regions. *Celestial Mechanics and Dynamical Astronomy*, 97(4):289–304, 2007. doi: https://doi.org/10.1007/s10569-007-9065-x.

D. A. Vallado. *Fundamentals of Astrodynamics and Applications*. Microcosm Press, 5th edition, 2022. ISBN 978-1881883227.

J. L. Worthy and M. J. Holzinger. Incorporating uncertainty in admissible regions for uncorrelated detections. *Journal of Guidance, Control, and Dynamics*, 38(9):1673–1689, 2015. doi: https://doi.org/10.2514/1.G000890.

E. M. Zucchelli, E. D. Delande, B. A. Jones, and M. K. Jah. Multi-fidelity orbit determination with systematic errors. *The Journal of The Astronautical Sciences*, 68:695–727, 2021. doi: https://doi.org/10.1007/s40295-021-00267-y.