5-2023

High Efficiency Angles-Only Space-Based Approaches for Geosynchronous Orbit Catalog Maintenance With Sparse Information

Louis M. Tonc
Utah State University

HIGH EFFICIENCY ANGLES-ONLY SPACE-BASED APPROACHES FOR GEOSYNCHRONOUS ORBIT CATALOG MAINTENANCE WITH SPARSE INFORMATION

by

Louis M. Tonc

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Aerospace Engineering

Approved:

Geordie Richards, Ph.D. Major Professor	David Geller, Ph.D. Committee Member
Stephen Whitmore, Ph.D. Committee Member	Douglas Hunsaker, Ph.D. Committee Member
Charles Swenson, Ph.D. Committee Member	D. Richard Cutler, Ph.D. Vice Provost for Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2023

ABSTRACT

High Efficiency Angles-Only Space-Based Approaches for Geosynchronous Orbit Catalog Maintenance with Sparse Information

by

Louis M. Tonc, Doctor of Philosophy

Utah State University, 2023

Major Professor: Geordie Richards, Ph.D.
Department: Mechanical and Aerospace Engineering

The thesis of this dissertation proposes a novel filter algorithm to improve tracking and catalog maintenance of uncooperative satellites and other Resident Space Objects (RSOs) in Geosynchronous Equatorial Orbit (GEO). Tracking can be supported by space-based tracking from observer satellites (OBSs). Practical limitations can lead to long time gaps between measurement updates when tracking RSOs from an OBS, which may induce a loss of fidelity or divergence of the estimation algorithm. The Extended Kalman filter (EKF) is commonly used for tracking RSOs but it diverges as a consequence of nonlinearity in the dynamics and nonlinearity in the optical measurements from OBSs. Both nonlinearities cause the underlying probability density (PDF) of the state vector to deviate from a Gaussian distribution, leading to divergence after measurement update. The Unscented Kalman filter (UKF) and the Gaussian Mixture Model filter (GMM) were proposed to solve the divergence problem in the EKF. A hybrid algorithm, the Hybrid Kalman-particle filter (HKF), was developed and likewise assessed to improve on the EKF methodology by combining with particle filtering techniques. Lastly, this work presents a novel filter algorithm, the Extended Step-Back Kalman filter (ESBKF), in which the measurement update is applied at a time in the past when the distribution of the RSO in state-space is approximately

Gaussian. The filter statistics are then propagated forward to the present, and the nonlinear effects of the dynamics are dramatically reduced, thereby avoiding divergence longer. In this document, the mathematical structure and properties of the aforementioned estimation algorithms and their utility are demonstrated in tracking simulations of an RSO in a GEO orbit being updated from four different scenarios of OBS availability.

(182 pages)

PUBLIC ABSTRACT

High Efficiency Angles-Only Space-Based Approaches for Geosynchronous Orbit Catalog Maintenance with Sparse Information

Louis M. Tonc

The thesis of this dissertation proposes a novel filter algorithm to improve tracking and catalog maintenance of uncooperative satellites and other Resident Space Objects (RSOs) in Geosynchronous Equatorial Orbit (GEO). Tracking can be supported by space-based tracking from observer satellites (OBSs). Practical limitations can lead to long time gaps between measurement updates when tracking RSOs from an OBS, which may induce a loss of fidelity or divergence of the estimation algorithm. The Extended Kalman filter (EKF) is commonly used for tracking RSOs but it diverges as a consequence of nonlinearity in the dynamics and nonlinearity in the optical measurements from OBSs. Both nonlinearities cause the underlying probability density (PDF) of the state vector to deviate from a Gaussian distribution, leading to divergence after measurement update. The Unscented Kalman filter (UKF) and the Gaussian Mixture Model filter (GMM) were proposed to solve the divergence problem in the EKF. A hybrid algorithm, the Hybrid Kalman-particle filter (HKF), was developed and likewise assessed to improve on the EKF methodology by combining with particle filtering techniques. Lastly, this work presents a novel filter algorithm, the Extended Step-Back Kalman filter (ESBKF), in which the measurement update is applied at a time in the past when the distribution of the RSO in state-space is approximately Gaussian. The filter statistics are then propagated forward to the present, and the nonlinear effects of the dynamics are dramatically reduced, thereby avoiding divergence longer.

ACKNOWLEDGMENTS

I would like to acknowledge those who have made this research effort possible. Their support and advisement have been invaluable.

First, I wish to acknowledge my advisor, Dr. Geordie Richards. His support and confidence in me made this possible. Dr. Geordie Richards provided me with excellent guidance and mentoring through all the hurdles to complete this project.

Second, I wish to acknowledge Dr. David Geller whose expert advice and encouragement of this project have made this possible. The knowledge I gained from Dr. David Geller in the fields of astrodynamics, optimal estimation, Monte Carlo analysis, and much more were essential in completing this project.

Third, I wish thank the Utah NASA Space Grant Consortium who has provided financial support for three years of this research project. Their support has made this work become a reality.

Lastly, none of this would have been possible without the love and support from my family. They have supported me from the beginning to pursue this degree.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Accurate tracking and estimation of RSOs in GEO pose a challenge for future travel and satellite operations. Obstacle avoidance necessitates real-time knowledge of the statistics of the target state, specifically the position and velocity vectors of the RSO with high probability. This requires the use of an orbit estimation algorithm that is capable of handling sparse observations while maintaining computational efficiency [1]. The Extended Kalman filter is the most commonly used estimation algorithm. However, there are scenarios in which the EKF has been observed to fail under reasonable conditions [2]. The effectiveness of the EKF is dependent on the duration between measurements.

Based on the literature, the EKF has been known to have difficulties in estimating the mean and covariance of the target state in nonlinear settings [3]. Objects in orbit are subject to nonlinear dynamics and are often updated by nonlinear measurements, as is the case with measurements based on optical angles. In this study, two sources of failure were identified for the EKF in nonlinear settings. The primary source is that the PDF that describes the relative likelihood of the RSO, represented by the combination of the position and velocity vectors into a single state vector, initially begins as a defined Gaussian distribution, but as time progresses, the distribution becomes skewed. The nonlinear orbital dynamics force the state vector to follow an orbital trajectory, which causes the PDF to become increasingly skewed due to the development of curvature. In a Cartesian coordinate system, the PDF will eventually no longer be adequately approximated as a Gaussian distribution.

Over the last several decades, several alternative filtering algorithms have been developed to resolve nonlinear effects. One such class of filters is known as particle filters, which use nodes or particles generated from an initial Gaussian distribution and propagate the particles through the nonlinear dynamics to obtain posterior mean and covariance values via sample statistics. A subset of particle filters are sigma-point filters, with on such filter be-

ing the UKF. The UKF creates sigma points which are particles obtained deterministically based on the matrix square root of the covariance matrix. Associated weights are assigned to each sigma point particle and are then propagated through the dynamical model. The deterministic used to obtain the sigma points allows for obtaining second-order accurate statistics without the need for orders of magnitude more particles [4].

The UKF can however break down when the target state density is severely skewed prior to performing a measurement update. Nevertheless, it yields a larger post-measurement covariance which compensates for bias in the mean. Another common algorithm is the Gaussian Sum filter or Gaussian Mixture Model filter, which approximates the state PDF by a sum of weighted Gaussian distributions [5]. It obtains a better estimate of the posterior mean and covariance since the geometry of the state PDF, which has curvature, is approximated by multiple Gaussian distributions in specific locations. However, this algorithm is computationally expensive and also requires updating the weights of each particle by evaluating various Gaussian distributions, which can be numerically sensitive.

Another filtering algorithm investigated is the Closed Skew-Normal Unscented Kalman filter (CSNUKF). This algorithm is an unscented form of the Closed Skew-Normal Kalman filter (CSNKF), which has propagation and measurement methodology similar to the EKF but with different parameters being tracked rather than the mean and covariance. The parameters being tracked define what is referred to as a closed skew-normal distribution, which is a PDF defined by the product of a Gaussian distribution and a Normal cumulative distribution function (CDF) which yields a Gaussian distribution with skewness. It was initially anticipated that the CSNUKF would generate a PDF with the appropriate skewness during the propagation phase, but this was not the case. Furthermore, the performance evaluation was extremely computationally burdensome since a numerical sampling technique, in this case a Markov chain Monte Carlo (MCMC) algorithm, was required to obtain the mean and covariance statistics.

This research effort presents two novel algorithms: the HKF and ESBKF. The HKF is a fusion of the EKF with a particle filter. Using a particle-filter strategy allows for using

a more accurate mean value and covariance matrix when performing the Kalman update. This algorithm is a preliminary approach that was investigated before implementing other more complex filter algorithms. This relatively straight forward approach was examined to see if a simpler remedy could be used. As demonstrated in the RESULTS chapter of this thesis, the performance capability of the HKF to handle divergence was effective in most conditions.

The ESBKF, however, was found to make improvements over the other filter algorithms by eliminating concerns of approximating the non-Gaussian PDF of the target state during the propagation phase. The updated mean was closer to the actual location of the RSO with the altitude being more accurate, in particular. The primary foundation of the ESBKF based on applying a measurement update to the last point in time in the past when the PDF was either exactly Gaussian or well-approximated as a Gaussian. This removes the nonlinear effects of dynamics, leaving only the measurement model with nonlinear effects. Once the mean and covariance are updated, they are propagated to the present point in time and the algorithm continues in a recursive manner. The ESBKF has a computational cost very close to the EKF.

Monte Carlo analyses were conducted on all of the filter algorithms and generated sample statistics using two different approaches. There were three performance metrics for the filters: 3σ plots comparing standard deviations obtained from Monte Carlo analyses, plots of the overlapping index calculated from two different standard deviations, and probability density function scatter plots. For technical details on the format of these plots used to evaluate filter performance can be found in the Details of Plots section in the RESULTS chapter.

A more robust closed-form analysis was attempted to yield a mathematically deductive performance metric for the ESBKF, but due to nonlinearities complicating a closed-form solution to integrate within a Gaussian distribution, this was not feasible. Since the ESBKF is a novel algorithm developed in this research effort, it was incumbent to match the simulation results with a developed mathematical theory. This is a topic for future work.

1.1 Dissertation Thesis Statement

The thesis of this dissertation is that divergence of the Extended Kalman filter, when tracking RSOs in GEO using right-ascension and declination angle measurements, can be resolved with only modest increases in computation time using an Extended Step-Back Kalman filter.

1.2 Dissertation Overview

Chapter 2 details a comprehensive literature survey of the background information related to the physical environment, as well as algorithms and techniques to be used in simulations. Chapter 3 outlines the mathematical details of the simulation, such as the dynamics and measurements, as well as the technical formulations of the filtering algorithms. Chapter 4 presents the results of the simulations along with corresponding analyses. Finally, Chapter 5 provides a summary of the work performed, along with suggested areas for future research.

1.2.1 Research Objectives and Corresponding Tasks

The following list details the objectives this research followed along with corresponding tasks. Main bullet points are the objectives and sub-bullet points are the tasks associated with each objective.

1. Identify the source of EKF divergence.
    a. Develop simulation of an RSO in GEO tracked by an OBS.
    b. Use simulation to isolate source of EKF divergence.
2. Assess performance of known alternative filter algorithms for resolving divergence of the EKF.
    a. UKF - Model four scenarios of RSO and OBS trajectories
    b. GMM - Model four scenarios of RSO and OBS trajectories.
3. Define a novel filter algorithm (ESBKF), and establish its theoretical properties.

   a. Define the propagation and update procedure.

   b. Prove that the ESBKF is identical to EKF under linear conditions.

4. Assess performance of ESBKF and compare with other algorithms.

   a. Model four scenarios of RSO and OBS trajectories.

   b. Demonstrate superior ability of ESBKF to stay within the bounds of overlapping index limit in each modeled scenario.

   c. Identify computational efficiency of ESBKF and compare with other algorithms.

CHAPTER 2
LITERATURE SURVEY

2.1 Geosynchronous Orbit

There is a certain orbital radius around Earth at which the angular rate of a circular orbit will match the rotation rate of Earth about its axis. This orbit is referred to as a Geosynchronous orbit (GEO). The main utility of this orbit is for communication, since the orbiting satellite will remain fixed above the ground relative to an observer on the surface of the Earth. If the satellite or resident space object (RSO) in GEO has a large inclination, the object relative to an observer on Earth will move up and down along a path perpendicular to the horizon. According to the European Space Agency (ESA), their database known as Database and Information System Characterising Objects in Space (DISCOS) details that the vast majority have a low inclination [2]. Therefore, the GEO objects simulated in this research were selected to have a nearly zero inclination.

Uncooperative satellites are spacecraft which do not have two-way signaling capability with whomever is tracking the object. This limits the methods for measuring these spacecraft, for example Doppler range-rate measurements typically require two-way communication. Optical measurements from either the ground or an observer spacecraft can be used for tracking uncooperative spacecraft as well as other orbital debris.

2.2 Ground-Based Angle Measurements for Tracking

Ground-based Electro-Optical Deep Space Surveillance (GEODSS) is a ground-based optical telescope system operated by the Air Force Space Command (AFSPC), which catalogs RSOs for multiple reasons. There are three global sites that comprise the locations of the system. This system tracks an RSO by collecting data in the electro-optical visible spectrum and therefore generates a line-of-sight vector. This method is much cheaper compared

to a RADAR system, which would give additional information for the range [6]. There are many governmental agencies worldwide invested in space surveillance tracking systems [7]; however, since the number of objects in orbit around Earth is always increasing, the need to have alternative means for tracking is desirable.

2.3 Space-Based Angle Measurements for Tracking

The primary means of measuring space objects, particularly uncooperative objects, is angle measurements, which yield a line of sight from the observer to the object being tracked via right ascension and declination angles. These two angles are then applied to a filtering algorithm, which generates a new estimate for the position and velocity vectors. Ground-based telescopes have typically been used to acquire angles-only measurements; however, there are obstacles that can limit the feasibility of this tracking method. These obstacles include weather interference, inflexibility of measurement locations, and lighting conditions.

The feasibility of space-based optical measurements is validated by the Midcourse Space Experiment (MSX) satellite [8], which had a Space-Based Visible (SBV) sensor using a telescope along with an electronics package used to look for visible streaks indicating the presence of an RSO. The MSX satellite was used by the United States Air Force to improve space-based surveillance for ballistic missile defense capabilities, and this was conducted in 2000.

2.4 Orbit Estimation Methodology

The objective of tracking objects in space is to catalog their orbits. Orbits are defined by the position and velocity vectors, that are combined to form a state vector and tracking of this vector through time is accomplished using filtering algorithms. The error in orbit estimation is also tracked since a perfect deterministic model for orbit propagation cannot be achieved. Error in orbit estimation is modeled as the covariance matrix. Filtering algorithms use stochastic modeling rather than purely deterministic modeling. Perturbing acceleration values that can be modeled by deterministic means in the orbital dynamics are gravitational

effects from the Earth not being a homogeneous sphere, the sun, the moon, and other planets in the solar system as well as solar radiation pressure. Not all perturbing accelerations can be modeled accurately and thus stochastic modeling will have to be applied. Nonetheless, modeling of orbital dynamics around the Earth is well-modeled deterministically and the time-varying stochastic components are much smaller in magnitude [9]. Depending on the filtering algorithm used, the process noise can be modeled stochastically by including a matrix term in the differential equations used to propagate the covariance matrix or by adding a matrix directly to the covariance matrix at discrete time steps, which represents the error in the state vector explicitly from the process noise over a single time step. Following sections in this document lay out how the process noise is applied in stochastic modeling approach. Measurements for the right ascension and declination angles are also not perfect, and the measurement error is modeled as Gaussian white noise. Similar to the process noise, implementation of the measurement noise is applied in different approaches, depending on the filtering algorithm.

2.5 Kalman Filter

As described by Maybeck [10], the Kalman filter is an optimal state estimation or tracking algorithm when the dynamics for the state vector and the measurement update model are both linear and have respective underlying probability density functions that are Gaussian. Initiation of the Kalman filter begins with defining a state vector with a defined mean value and covariance matrix. A mean value and covariance can exclusively represent a Gaussian distribution, which is the assumed PDF. The dynamical system, being linear in nature, means that the class of PDF will not be altered as it is propagated through time. In other words, the PDF will remain a Gaussian albeit with a new mean and covariance to define the Gaussian distribution. When a measurement is available to update the state vector, the measurement PDF in the state space domain will be Gaussian because the PDF in the measurement domain is defined as a Gaussian and since the measurement model is linear with respect to the state, the PDF mapped into the state space domain will be Gaussian. The Kalman filter combines the state vector PDF with the measurement PDF

and acquires a new Gaussian distribution. The Kalman filter is optimal because the updated PDF incorporates all available measurements and knowledge of measurement devices, as well as prior knowledge of the system in such a way that error is minimized statistically [10]. This is represented by this new Gaussian distribution and the corresponding mean value is the optimal location in a Kalman filter.

2.5.1 Extended Kalman Filter

Most dynamics encountered in the real world are nonlinear as well as most measurements taken. This is the case for orbital dynamics with angles-only measurements. In order to apply the Kalman filter to nonlinear systems, the typical engineering approach is to linearize the dynamics and measurement model about a nominal value and apply the Kalman filter equations to this linearized approximation, as described by Woodburn et. al [11]. This approach when applied to the Kalman filter is called the Extended Kalman filter (EKF) and is likely the most widely used approach when estimating nonlinear systems [12]. The nominal value about which the Jacobians, etc. are linearized is the current filter estimate of the state vector. This linearized approach is advantageous in that it allows the use of the Kalman filter equations however, depending on the dynamics and measurement model, the calculations of the Jacobians may be computationally burdensome.

2.6 Alternative Filters to the EKF

There has been great interest in developing filters that do not suffer from limitations due to nonlinear effects and can accurately estimate the state vectors. As a result, various alternative filters have been proposed using various different methodologies. These methodologies include propagating multiple samples of the state vectors to obtain sample statistics, applying multiple EKF samples with corresponding weights to obtain weighted sample statistics, using a skewed PDF rather than a Gaussian, and backwards smoothing the solution iteratively. Some methods require a higher computational demand but increase accuracy, while other methods could be used as a batch-processing, where computational time is not as important since the algorithm would not be used in real-time.

2.6.1 Particle Filters

Mashiku et al. introduced a general particle filter as a method to obtain the statistics of the state vector regardless of nonlinearities. This class of filters uses multiple samples of the state vector, or particles, which collectively run through a simulation, and the resulting sample statistics are used as the means of performance evaluation [13]. Each particle is assigned a weighting value such that when a measurement is received, the weighting values will be altered based on the evaluation of the measurement PDF.

There is an issue described by Pardal et al. with what is referred to as "sample impoverishment" [14], where after a measurement is received and processed, there are a large number, perhaps the vast majority, of particles whose weight contribution essentially goes to zero and therefore no longer contribute valuable data for sample statistics. Carrying particles that are useless for filter knowledge is an unnecessary waste of computational efforts. Since a large number of particles are typically required to avoid sample impoverishment, the generic particle filter is not computationally feasible. Fortunately, there are variations of the particle filter with more structure which improve computational efficiency.

2.6.2 Unscented Kalman Filter

The linearized approximation used in the EKF may not be adequate to handle nonlinearities in certain scenarios. To address this concern, there is a class of particle filters referred to as Sigma Point filters of which a commonly used variant is the Unscented Kalman filter [15]. The UKF methodology aims to capture higher-order accuracy by propagating sigma point particles through the dynamics as well as applying measurement model evaluations to each particle and then performing weighted statistics. The UKF has been seen to have improvements over the EKF in many fields of study [16], not only in orbital dynamics, as in this study. There are variations of the UKF, such as the Square Root Unscented Kalman filter (SRUKF). The SRUKF improves on computational expediency by avoiding the use of Cholesky decomposition, which is used in obtaining the matrix square-root of the covariance matrix. The SRUKF propagates the matrix square-root directly by QR factor-

ization and Cholesky updating [17]. While better accuracy of the mean and covariance is achieved by the UKF compared to the EKF there is still the issue that the true PDF is not simply a Gaussian distribution. The UKF only refines the mean and covariance, but does not track a PDF that has skewness in its geometry.

2.6.3 Iterative Extended Kalman Filter

In order to reduce the errors in the EKF due to nonlinearities, specifically in the measurement model, an iterative version of the EKF was developed, referred to as the Iterative Extended Kalman Filter (IEKF). This algorithm re-linearizes the measurement model about the update state vector immediately after a measurement update, and applies the Kalman update procedure recursively until convergence. There is a Gauss-Newton search for optimization to maximize the posterior estimate [18].

2.6.4 Reverse and Backward-Smoothing Extended Kalman Filter

Several filtering algorithms have been developed which operate by propagating backwards in time after a measurement is received back to the previous point in time when a measurement was processed. These algorithms typically use a smoothing technique as well. One algorithm described and referred to as the Reverse Kalman filter by Taplin [19] essentially applies an optimal smoothing approach in which the forward and reverse algorithms are iterated until convergence to a common solution is achieved. This algorithm works by propagating the mean value and the covariance matrix backwards in time using equations derived specifically for this process and acquiring new estimates at the previous time when a measurement was received. The EKF methodogoly is then used again to propagate forward in time. This back and forth continues until a convergent result is attained, as described by Maybeck [10]. Reversing the backwards propagation uses a set of equations which propagate differently that what the Kalman filter implements in a forward propagation approach, which are detailed by Maybeck. However, since initial conditions are the only known parameters at the onset of using a filtering algorithm and not a future value after a set of measurements are applied, the equations for propagating and implementing

measurements in reverse time cannot be applied at the onset. This is the reason why initially the traditional Kalman is used to achieve a forward estimate and then the reverse Kalman filter methodolgy can be applied. However, using this filtering approach was not pursued because when nonlinear effects yield errors when using the EKF in certain scenarios, using this approach will only start the with a flawed estimate for the reverse optimal smoothing.

A second algorithm is referred to as the Backward-Smoothing Extended Kalman filter. This algorithm is quite similar to the Reverse Kalman filter in that a smoothing scheme is performed in a forward and backward manner [20]. The primary difference is the optimization using a Gauss-Newton search [18]. These algorithms are following a much different methodology that the novel filtering algorithm from this research effort, the ESBKF.

2.6.5 Gaussian Mixture Model Filter

The Gaussian Mixture Model filter (GMM) is a variant of a class of filters known as Ensemble Kalman filters, which in turn, are a subclass of the particle filters. Ensemble Kalman filters typically use particles with uniform weighting factors to obtain posterior distributions, similar to the general particle filter, but the unique addition is that each particle forms the center of a Gaussian kernel. An Ensemble Kalman filter is essentially a large number of individual particles with each individually following the EKF methodology until a measurement is obtained and then weighting factor values are updated following particle filter methodology but in a more robust manner as described by Stordal et. al [21]. The GMM has a unique feature to the Ensemble Kalman filter class in which the particles have weighting factors that are more robustly obtained as well as deterministic locations and parameters for the Gaussian kernels. Each particle initially is assigned a location, weighting value, and covariance matrix through an algorithm known as Gauss-Hermite quadrature [22]. After all the particles are assigned initial conditions, they sequentially propagate forward to the next time step at which point either the weighted sample statistics are immediately calculated and stored or a measurement is processed and weight values are updated and then sample statistics are evaluated. The primary interest in this algorithm is that it can better approximate the state vector PDF region of high probability since the

GMM particles will propagate into locations of high probability with respect to the true PDF and then since each particle has an associated Gaussian kernel, the collective mixture of all the Gaussian kernels will fill in the regions of high probability. As with particle filter in general, the large number of particles required does reduce the computational efficiency.

The GMM is unique among the Ensemble Kalman filter class in that the particles have weighting factors that are more robustly obtained as well as deterministic locations and parameters for the Gaussian kernels. Each particle is initially assigned a location, weighting value, and covariance matrix through an algorithm known as Gauss-Hermite quadrature [22]. Once all the particles are assigned initial conditions, they sequentially propagate forward to the next time step. At which point, either the weighted sample statistics are immediately calculated and stored or a measurement is processed, the weight values are updated, and then sample statistics are evaluated.

The primary advantage of the GMM is that it can better approximate region of high probability of the state vector PDF. This is a result of the GMM particles propagating into locations of high probability with respect to the true PDF and then since each particle has an associated Gaussian kernel, the collective mixture of all the Gaussian kernels will fill in the regions of high probability. As with particle filter in general, the large number of particles required does reduce the computational efficiency.

2.6.6 Closed Skew-Normal Kalman Filter

The GMM algorithm can approximate the true PDF in the future, but it requires a significant number of particles to do so. An alternative that is comparable to the EKF in terms of computational speed is desired. Looking at the geometry of the true PDF in the orbital plane, with respect to both position and velocity, shows skewness in a bending shape. Essentially, what would have been an elliptical geometry becomes bent along the orbital trajectory. Therefore, this research investigated an algorithm that uses a filter PDF with built-in parameters to yield a skewed geometry. This algorithm is the Closed Skew-Normal Kalman filter (CSNKF). As laid out by Naveau et al. [23], it tracks several parameters, and none of them explicitly represent the mean value or covariance matrix, except in the case

of zero skewness. These parameters allow shaping the PDF geometry to have more or less skewness and in what direction. Its computational speed is on par with the EKF. However, evaluating the performance metrics via the mean and covariance requires either numerical integration or random sampling statistics. There is an issue with random sampling in that the PDF from which samples are drawn is a closed skew-normal distribution, and there is no explicit way to draw samples. However, there is a class of methods called Markov chain Monte Carlo (MCMC) methods that allow a brute-force iterative technique to sample from any desired probability density function by evaluating and comparing states in the chain [24].

2.7 Metropolis-Hastings Algorithm

When generating random samples from a PDF which is not readily able to be drawn from, such as the case with a closed skew-normal PDF, an MCMC method can be used. The Metropolis-Hastings algorithm is a common choice for generating the Markov Chain in the MCMC algorithm [25]. Metropolis-Hastings essentially draws samples from any desired PDF as detail in the following algorithm sequence, seen in Table 2.1.

Table 2.1: Metropolis Hastings Algorithm

Initialize a starting value for random number, $\mathbf{X}$
Generate a random number, $\mathbf{Y}$, offset from $\mathbf{X}$ using a well-known random generator
Calculate the conditional probability of current value $\mathbf{X}$ given candidate $\mathbf{Y}$, $q(\mathbf{X}\|\mathbf{Y})$
Calculate the conditional probability of candidate $\mathbf{Y}$ given current value $\mathbf{X}$, $q(\mathbf{Y}\|\mathbf{X})$
Calculate probability of $\mathbf{X}$ from PDF desired to sample from, $p(\mathbf{X})$
Calculate probability of $\mathbf{Y}$ from PDF desired to sample from, $p(\mathbf{Y})$
Let $\beta = \min\left(\frac{p(\mathbf{Y})q(\mathbf{X}\|\mathbf{Y})}{p(\mathbf{X})q(\mathbf{Y}\|\mathbf{X})}, 1\right)$
Generate a random number, α, from uniform distribution between 0 and 1
If α is less than or equal to β, then candidate $\mathbf{Y}$ is accepted and replaces $\mathbf{X}$
This process is iterated, for example 500 runs, until the sample converges to a random draw from p

2.8 L_1 and L_2-Norm and Gaussian Distribution Comparisons

In the context of parameter estimation, it is desired to obtain the most likely values according to probabilities derived from measurements. Objective functions are used for minimizing errors to obtain the most probable solution. Two examples of objective functions for parameter estimation are the L_1 and L_2-norm. The L_2-norm is also known as the Least-Squares Method (LSM) and is widely used since calculations are easy to perform, and it yields a unique solution [26] that is also optimal if the errors are Gaussian distributed [27]. This method, however, assumes that the measurements have a normal distribution, and if there are gross or systematic errors in the measurements, then the L_2-norm will not yield as good an estimate. When these errors are present, the L_1-norm can be used to give a better result. This method, also known as the Least-Absolute-Values Method (LAVM), handles gross errors better but lacks computational simplicity when compared to the LSM approach. An iterative, linear programming approach, such as the Simplex Method, is often required to solve the minimization problem using the LAVM. There are also usually more than one solution, whereas the LSM has a unique solution.

A practical application of these objective functions used in this research effort is when approximating a single Gaussian distribution by a weighted-sum of multiple Gaussian distributions, referred to as a Gaussian mixture. The selection of a covariance scaling parameter, k, is made using an L_2-norm criterion [5].

This study will present simulation results in a quantitative and qualitative manner. A Monte Carlo simulation will generate errors between the estimate of the state of the RSO and the true state as well as a covariance matrix by the filter algorithm. The first involves calculating the sample standard deviation of mean value errors. The second involves calculating the sample mean of standard deviations, taken from the square root of the diagonal elements of the filter covariance matrix. If the filter algorithm is adequately tracking the RSO, these two standard deviations should be close.

In comparing two univariate normal distributions, the L_1-norm was used to calculate a quantity of overlap between the two PDFs, which is referred to as the overlapping index,

η. A threshold of 64% for the overlapping index was based on an example from Pastore et al. [28] in which this threshold was said to demonstrate the two probability densities are not so different. All six state vector elements with have their respective η values on a single plot per simulation. Each simulation will have a specific OBS configuration and a specific time of measurement. For each time of measurement, four subplots representing the four OBS configurations will be combined into a single plot.

2.9 Encke's Method

When two-body Keplerian dynamics are used to model the orbit of an RSO or OBS, the solution does not require numerical integration of the position and velocity vectors since there is a semi-analytic closed-form solution. However, when using a dynamical model that more accurately accounts for real-world perturbing accelerations, such as gravitational zonal harmonics and lunar and solar third bodies, numerical integration will be required. The direct method of simply adding the perturbing accelerations to the two-body gravitational acceleration and integrating the position and velocity vectors forward in time, typically using a Runge-Kutta scheme, is known as Cowell's Method [29]. There is a disadvantage of using this straightforward approach which is the time step can't be too large otherwise the numerical integration will become exceedingly inaccurate. There is another approach which utilizes that fact that the two-body gravitational acceleration is orders of magnitude larger than the perturbing accelerations and therefore a Keplerian orbit will track closely with the true orbital trajectory for a reasonably long duration. If instead of attempting to propagate the true orbital trajectory by integrating the position and velocity directly, one were to integrate the deviations of both the position and velocity vectors from those of a referenced Keplerian orbit, the time step for numerical integration can be increased significantly while maintaining accuracy. This method is called Encke's Method [29]. This was the approach used in this research effort for all of the filtering algorithms and also when generating the truth models for the RSO and OBS.

CHAPTER 3
PROBLEM FORMULATION AND SETUP

3.1 Background

Tracking and surveillance of uncooperative satellites and other RSOs, in GEO is vital for ensuring the continuation of communication, weather forecasting, and other space-based operations. Ground-based tracking is commonly implemented to estimate the position and velocity of objects in GEO, however, weather interference and limitations due to size and range of RSOs can reduce the effectiveness of this approach. Incorporating space-based tracking using an OBS, which acquires angle measurements in the form of right-ascension and declination, can ameliorate these problems and improve on the performance of ground-based observations. Due to the vast quantity of RSOs that lie in GEO, it is not economically feasible to dedicate an observer spacecraft for each RSO to update the position and velocity in a frequent manner. Correspondingly, there can be long time gaps between measurement updates when tracking RSOs, which may induce a loss of fidelity or divergence of the estimation algorithm. This has been seen in simulations when using the EKF for tracking. This divergence has been identified as a consequence of nonlinearity in the dynamics and nonlinearity in the optical measurements, both of which cause the underlying probability density of the state vector to deviate from a Gaussian distribution. Figure 3.1 depicts an example of an EKF undergoing divergence after a measurement update. See Details of Plots section in the RESULTS chapter for an explanation of what Figure 3.1 is illustrating.

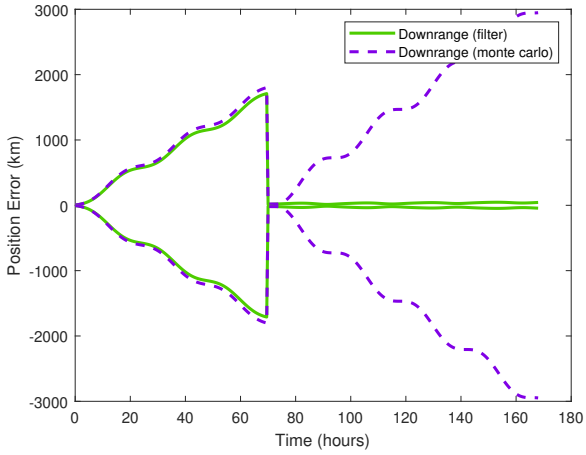


Fig. 3.1: Extended Kalman filter divergence example

3.2 Formulation

A computer simulation was created which incorporated a generic RSO in GEO, and an OBS. Each OBS has a space-based sensor providing a line of sight measurement to the RSO via two angle measurements, the right-ascension and declination. Filtering algorithms were run through the simulation to determine performance of each. Four different scenarios of the observer satellite orbital trajectory were investigated to see if the physical location of the observer satellite could be a simple remedy to the divergence of the EKF. They were a medium altitude, high altitude, high inclination, and lastly a combination of the three. A novel filtering algorithm, the ESBKF, has been developed in this research effort and applied to the simulation. It has the advantage over particle filters by being computationally expedient near the order of the Extended Kalman filter, while maintaining accuracy. The performance of the SBKF will be compared against the other filtering algorithms which are the Unscented Kalman filter, Hybrid Kalman filter, Gaussian Mixture Model filter, and the Unscented Closed Skew-normal Kalman filter. Initial conditions and parameters of the simulation were similar to those based on efforts by Andrews [2]. The dynamical model for all the filters was a Keplerian two-body orbital system, however, the RSO truth model involved dynamics taking into consideration a higher gravity model as well as lunar and solar effects.

3.3 Dynamical System

The dynamical system for modeling the true trajectory of the RSO involves a gravity model incorporating up to J4 gravitational zonal harmonics along with the point mass of Earth. In addition, the solar and lunar third body perturbing gravitational accelerations and additive process noise simulated as Gaussian white noise were also included. It was found that applying the perturbing accelerations in the filter model did not improve filter accuracy significantly to justify the computational burden of numerical integration and therefore the much simpler Keplerian orbital dynamics were used in the estimation algorithms.

The position vector and velocity vectors of the RSO are defined respectively as $\mathbf{r}$ and $\mathbf{v}$. Likewise, the position and velocity vectors for the observer satellite are defined as $\mathbf{r}_{obs}$ and $\mathbf{v}_{obs}$, respectively. Since the equations of motion are based on second-order differential equations, a restructuring needs to take place to use the first-order numerical integration scheme, Runge-Kutta integration. In the state space formulation, the state vector for the RSO is $\mathbf{X}$ and the state vector for the observer satellite is $\mathbf{X}_{obs}$. This is represented as,

$$\mathbf{X} = \begin{bmatrix} \mathbf{r} \\ \mathbf{v} \end{bmatrix}, \quad \mathbf{X}_{obs} = \begin{bmatrix} \mathbf{r}_{obs} \\ \mathbf{v}_{obs} \end{bmatrix}. \tag{3.1}$$

Within the state space regime, the equations of motion for the RSO and OBS are $\dot{\mathbf{X}} = \mathbf{L}(\mathbf{X})$ and $\dot{\mathbf{X}}_{obs} = \mathbf{L}(\mathbf{X}_{obs})$ where $\mathbf{L}$ is calculated by the following,

$$\mathbf{L}(\mathbf{X}) = \begin{bmatrix} \mathbf{v} \\ -\mu\left(\frac{\mathbf{r}}{\|\mathbf{r}\|^3}\right) + \mathbf{a}_{J2} + \mathbf{a}_{J3} + \mathbf{a}_{J4} + \mathbf{a}_{solar} + \mathbf{a}_{lunar} + \mathbf{w} \end{bmatrix}. \tag{3.2}$$

The gravitational zonal harmonics perturbing accelerations require the value of the Earth's equatorial radius R_{eq} and empirically evaluated coefficients J_2, J_3, and J_4 and are calculated as follows,

$$\mathbf{n} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T, \quad \mathbf{i}_r = \frac{\mathbf{r}}{|\mathbf{r}\|}$$

$$
\begin{aligned}
A &= -\frac{\mu R_{eq}^2 J_2}{2\|\mathbf{r}\|^3}, \\
B &= 3\left(\mathbf{n}^T\mathbf{i}_r\right)^2 - 1, \\
\mathbf{A}_r &= -\frac{\mu R_{eq}^2 J_2}{2}\left(-\frac{3}{\|\mathbf{r}\|^4}\right)\mathbf{i}_r, \\
\mathbf{B}_r &= 6\left(\mathbf{n}^T\mathbf{i}_r\right)\left(\left(\frac{1}{\|\mathbf{r}\|}\right)\left(I_{3\times 3} - \mathbf{i}_r\mathbf{i}_r^T\right)\mathbf{n}\right), \\
\mathbf{a}_{J2} &= A\mathbf{B}_r + B\mathbf{A}_r,
\end{aligned}
\tag{3.3}
$$

$$
\begin{aligned}
A &= -\frac{\mu R_{eq}^3 J_3}{2\|\mathbf{r}\|^4}, \\
B &= 5\left(\mathbf{n}^T\mathbf{i}_r\right)^3 - 3\left(\mathbf{n}^T\mathbf{i}_r\right), \\
\mathbf{A}_r &= -\frac{\mu R_{eq}^3 J_3}{2}\left(-\frac{4}{\|\mathbf{r}\|^5}\right)\mathbf{i}_r, \\
\mathbf{B}_r &= \left(15\left(\mathbf{n}^T\mathbf{i}_r\right)^2 - 3\right)\left(\left(\frac{1}{\|\mathbf{r}\|}\right)\left(I_{3\times 3} - \mathbf{i}_r\mathbf{i}_r^T\right)\mathbf{n}\right), \\
\mathbf{a}_{J3} &= A\mathbf{B}_r + B\mathbf{A}_r,
\end{aligned}
\tag{3.4}
$$

$$
\begin{aligned}
A &= -\frac{\mu R_{eq}^4 J_4}{8\|\mathbf{r}\|^5}, \\
B &= 35\left(\mathbf{n}^T\mathbf{i}_r\right)^4 - 30\left(\mathbf{n}^T\mathbf{i}_r\right)^2 + 3, \\
\mathbf{A}_r &= -\frac{\mu R_{eq}^4 J_4}{8}\left(-\frac{5}{\|\mathbf{r}\|^6}\right)\mathbf{i}_r, \\
\mathbf{B}_r &= \left(140\left(\mathbf{n}^T\mathbf{i}_r\right)^3 - 60\left(\mathbf{n}^T\mathbf{i}_r\right)\right)\left(\left(\frac{1}{\|\mathbf{r}\|}\right)\left(I_{3\times 3} - \mathbf{i}_r\mathbf{i}_r^T\right)\mathbf{n}\right), \\
\mathbf{a}_{J4} &= A\mathbf{B}_r + B\mathbf{A}_r,
\end{aligned}
\tag{3.5}
$$

Third-body perturbing accelerations both from the sun and the moon are calculated by the following,

$$\begin{aligned} \boldsymbol{\rho} &= \mathbf{r}_{body}, \\ \boldsymbol{\Delta} &= \boldsymbol{\rho} - \mathbf{r}, \\ \mathbf{a}_{body} &= \mu_{body} \left(\frac{\boldsymbol{\Delta}}{\|\boldsymbol{\Delta}\|^3} - \frac{\boldsymbol{\rho}}{\|\boldsymbol{\rho}\|^3} \right). \end{aligned} \tag{3.6}$$

The process noise term $\mathbf{w}$ is a three-dimensional zero-mean white noise process vector that is independent for the RSO and OBS and is described mathematically by the following,

$$\begin{aligned} \mathbf{w} &= \mathcal{N}\left(0_{3\times1}, Q_{3\times3}\right), \quad \mathbb{E}\left[\mathbf{w}\right] = 0_{3\times1}, \\ \mathbb{E}\left[\mathbf{w}\mathbf{w}^T\right] &= Q_{3\times3}\delta\left(t - t^{'}\right), \quad Q = \left(1.5 \cdot 10^{-10}\right)^2 I_{3\times3} \left(\frac{km^2}{s^2}\right) \end{aligned} \tag{3.7}$$

Dynamics modeled in filtering algorithms implement a Keplerian orbital model in which the state vector at a future point in time is calculated from a semi-analytic closed form solution. Not requiring numerical integration greatly improves computational expediency. The equations of motion $\mathbf{N}$ for the filters are the following,

$$\mathbf{N}(\mathbf{X}) = \begin{bmatrix} \mathbf{v} \\ -\mu\left(\frac{\mathbf{r}}{\|\mathbf{r}\|^3}\right) \end{bmatrix}. \tag{3.8}$$

Values of the parameters included in equations (3.1)-(3.8) are shown in Table 3.1.

Table 3.1: Dynamics Parameters

Parameter	Value
μ (km^3/s^2)	398600.436
R_{eq} (km)	6378.137
J_2	0.00108262583130715
J_3	-2.53261237062519 $\cdot$ 10^{-6}
J_4	-1.6161966 $\cdot$ 10^{-6}
μ_{moon} (km^3/s^2)	4902.7991086
μ_{sun} (km^3/s^2)	132712440017.987

It will be shown in the RESULTS chapter that **N** is used for the filters and not **L**. This was selected since **N** has a semi-analytical closed-form solution, in a Kepler routine, while **L** requires numerical integration which greatly increase computation time and there was no evident advantage of using the higher fidelity dynamical model.

3.4 Measurement Model

The position vector of the RSO centered about the observer satellite as represented in the inertial reference frame centered about the Earth is $\boldsymbol{\rho}$. The measurements taken by the observer satellite are two angles, the right ascension and declination. They are represented as α and δ, respectively. The measurement vector is the combination of the two angles represented by **Z**.

$$\mathbf{Z} = \begin{bmatrix} \alpha \\ \delta \end{bmatrix}.$$

$$\begin{aligned}
\boldsymbol{\rho} = \mathbf{r} - \mathbf{r}_{obs} &= \begin{bmatrix} \rho_x \\ \rho_y \\ \rho_z \end{bmatrix}, \quad \rho = \|\boldsymbol{\rho}\|, \\
\alpha = \tan^{-1}\left(\frac{\rho_y}{\rho_x}\right), &\quad \delta = \sin^{-1}\left(\frac{\rho_z}{\rho}\right).
\end{aligned} \tag{3.9}$$

The measurement Jacobian matrix H is defined by the following,

$$\begin{aligned}
\alpha_{\rho_x} &= -\frac{\rho_y}{\rho_x^2 + \rho_y^2}, \quad \alpha_{\rho_y} = \frac{\rho_x}{\rho_x^2 + \rho_y^2} \quad \alpha_{\rho_z} = 0 \\
\delta_{\rho_x} &= \left(\frac{1}{\sqrt{1 - \frac{\rho_z}{\rho^2}}}\right)\left(-\frac{\rho_x \rho_z}{\rho^3}\right), \quad \delta_{\rho_y} = \left(\frac{1}{\sqrt{1 - \frac{\rho_z}{\rho^2}}}\right)\left(-\frac{\rho_y \rho_z}{\rho^3}\right), \\
\delta_{\rho_z} &= \left(\frac{1}{\sqrt{1 - \frac{\rho_z}{\rho^2}}}\right)\left(\frac{1}{\rho} - \frac{\rho_z^2}{\rho^3}\right), \\
\frac{\partial \mathbf{Z}}{\partial \mathbf{r}} &= \left(\frac{\partial \mathbf{Z}}{\partial \boldsymbol{\rho}}\right)\left(\frac{\partial \boldsymbol{\rho}}{\partial \mathbf{r}}\right) = \begin{bmatrix} \alpha_{\rho_x} & \alpha_{\rho_y} & \alpha_{\rho_z} \\ \delta_{\rho_x} & \delta_{\rho_y} & \delta_{\rho_z} \end{bmatrix} \\
\frac{\partial \mathbf{Z}}{\partial \mathbf{v}} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
H = \frac{\partial \mathbf{Z}}{\partial \mathbf{X}} &= \begin{bmatrix} \frac{\partial \mathbf{Z}}{\partial \mathbf{r}} & \frac{\partial \mathbf{Z}}{\partial \mathbf{v}} \end{bmatrix}.
\end{aligned} \tag{3.10}$$

The measurement statistics are defined by a covariance matrix representing the accuracy of the right-ascension and declination angles as measured from the actual location of the RSO. The measurement statistics are shown below in Table 3.2.

Table 3.2: Angle Measurement Statistics

angle measurement type	standard deviation (rad)
right ascension	$5 \cdot 10^{-5}$
declination	$5 \cdot 10^{-5}$

The covariance matrix for the measurement, R, has the following value,

$$R = \left(5 \cdot 10^{-5}\right)^2 I_{2\times 2} \quad (rad^2) \tag{3.11}$$

3.5 Initial Conditions

The initial conditions for the RSO and the observer satellite are defined by the orbital elements shown in Table 3.4. For the RSO, this location is the mean value characterized by statistical estimation. The initial conditions used were slight deviations of those used by Andrews [2]. Andrews obtained the average orbital elements from a catalog of GEO objects to acquire the initial conditions for the RSO. Four case studies for the observer satellite were investigated with two of them following suit with Andrews, the third based on the potential of out-of-plane possibilities, and the last was a combination of all running collectively. The three cases for the observer satellite all have near-circular orbits. The first case or the base model has an altitude 7,342 km below GEO. The other two cases have the same initial conditions for the observer satellite except for one unique change for each. The second case has an altitude 1,000 km below GEO. Lastly, the third case has an inclination of 30 degrees. The second case was used to see if being closer to the RSO could possibly help with divergence issues. The third case was contrived based on the understanding that because the orbits of the OBS and RSO in the base model are nearly coplanar, the right ascension measurement is the most useful information in a measurement update since the declination is always near zero. By increasing the inclination to 30 degrees, the declination is therefore making a bigger contribution by giving an out-of-plane metric for the location of the RSO. It was hoped that this possibly may improve the measurement update step of the filtering algorithms by placing the measurement PDF more closely to the actual location of the RSO. The rationale is that since the skewness in the state PDF develops within the orbital plane and the line of sight vector between the OBS and RSO also lies nearly within the same plane, then the right ascension angle alone will not place the location of the RSO accurately when applying the EKF methodology. Since a line of sight measurement does not convey range information, the intersection of the measurement PDF and the state PDF

using the Gaussian assumption in the EKF will not be in the same location as would have been if the true state PDF were applied. However, for a line of sight that originates from an angle of incidence that is fairly large as would be the case from a 30 degree inclined OBS, the line of sight would lie in the vicinity of where the true state PDF would have placed it. Standard Deviation values for the RSO were established in the LVLH reference frame. The values are the following,

Table 3.3: RSO Initial Standard Deviations

LVLH Element	km or km/s
position altitude	10
position downrange	2
position crosstrack	2
velocity altitude	$1.45 \cdot 10^{-4}$
velocity downrange	$1.45 \cdot 10^{-4}$
velocity crosstrack	$7.27 \cdot 10^{-4}$

The modeling of when angle measurements were to be received and processed are dependent on the initial conditions for the observer satellite as shown in Table 3.4. It is the semi-major axis of the OBS that dictated the time of measurements to simulated. This is why 70 hours and 140 were used. As seen in the RESULTS chapter, the measurement time of 24 hours was also simulated. This time of measurement would not be physically possible, but it was artificially added to demonstrate the mathematical behavior of the filter algorithms when a measurement is take an earlier time.

Table 3.4: RSO and OBS Orbital Elements

orbital elements	**RSO**	**OBS 1**	**OBS 2**	**OBS 3**
semi-major axis (km)	42164	34822	41164	34822
eccentricity	10^{-5}	10^{-4}	10^{-4}	10^{-4}
inclination (deg)	0.5	1	1	30
right ascension (deg)	10	100	100	100
argument of perigee (deg)	240	120	120	120
true anomaly (deg)	130	220	220	220

3.6 Orbit Estimation Strategies

Filtering algorithms are used to estimate the state of an object in orbit. Traditionally, the Extended Kalman filter is the default algorithm implemented. In the event that the EKF breaks down, particle filters are often the backup approach used. There are pros and cons to all of the algorithms and is the primary motivation behind the thesis research efforts. The order of the filter algorithms presented below follows how they were simulated chronologically. This begins with the EKF which is known to fail under certain conditions thus being the catalyst for this research endeavor. Next, the HKF is investigated which takes the approach of combining the computational expediency of the Extended Kalman filter with the improved statistical accuracy of particle filters into a Hybrid algorithm. The UKF is the first introduction into particle filters as a means to extract the statistics induced from nonlinearity in the dynamics from sample values. The final algorithm was conceived out of efforts of this research project to somehow reduce or eliminate completely the nonlinear effects of the dynamics. This algorithm has been named the Step-Back Kalman filter. It essentially removes the problem of attempting to capture the true probability distribution underlying the state of the RSO having been propagated forward in time through Keplerian orbital dynamics. The Step-Back Kalman filter uses the mean value and covariance at a time prior to the measurement such that the probability distribution is well-represented as a Gaussian distribution.

3.6.1 Extended Kalman Filter

The Extended Kalman filter (EKF) is a nonlinear extension of the Kalman filter based on the premise that both the dynamics and measurement model are both linear systems [10]. The algorithm for the EKF, like all of the filters investigated in this research effort, has a propagation step and a measurement update step. These steps are designed to track the mean $\hat{\mathbf{X}}$ and covariance matrix P of the estimated state vector $\mathbf{X}$. The propagation step for the EKF is based on integrating the dynamical system (3.8), namely,

$$\hat{\mathbf{X}}(t^-) = \hat{\mathbf{X}}(t_0^-) + \int_{t_0}^{t} \mathbf{N}\left(\hat{\mathbf{X}}(\tau)\right) \mathrm{d}\tau. \tag{3.12}$$

The Jacobian of the dynamics (3.8) will be required to compute the state transition matrix, which in turn is used to propagate forward the covariance matrix of the state. The Jacobian of the dynamical system is computed from,

$$F\left(\hat{\mathbf{X}}(t)\right) = \left.\frac{\partial \mathbf{N}}{\partial \mathbf{X}}\right|_{\mathbf{X}=\hat{\mathbf{X}}(t)} = \begin{bmatrix} 0_{3\times 3} & I_{3\times 3} \\ -\left(\frac{\mu}{\|\mathbf{r}\|^3}\right)\left(I_{3\times 3} - 3\mathbf{i}_r\mathbf{i}_r^T\right) & 0_{3\times 3} \end{bmatrix}. \tag{3.13}$$

Indeed, the state transition matrix is obtained by means of numerical integration according to the formula

$$\Phi(t,t_0) = \left.\frac{\partial \mathbf{X}(t)}{\partial \mathbf{X}(t_0)}\right|_{\mathbf{X}(t_0)=\hat{\mathbf{X}}(t_0)} = I + \int_{t_0}^{t} F\left(\hat{\mathbf{X}}(\tau)\right)\Phi(\tau,t_0)\,\mathrm{d}\tau. \tag{3.14}$$

For more mathematical details on the source of this formula, see Appendix A.1. The state transition matrix is then used to propagate the covariance matrix P,

$$P(t^-) = \Phi(t,t_0)P(t_0^-)\Phi^T(t,t_0) + \int_{t_0}^{t} \Phi(\tau,t_0)GQG^T\Phi^T(\tau,t_0)\,\mathrm{d}\tau. \tag{3.15}$$

The last term in (3.15) involves Q which is the covariance of the process noise as defined in (3.7). The matrix G is a transformation matrix the maps Q into the state space covariance domain. Since the process noise is an acceleration vector the appropriate mapping of Q

will be in the lower right quadrant of the covariance since that is location that explicitly represents the velocity.

$$G = \begin{bmatrix} 0_{3\times3} \\ I_{3\times3} \end{bmatrix}. \tag{3.16}$$

Once a measurement $\tilde{Z}$ is available, the statistics are updated by

$$\begin{aligned} K(t) &= P(t^-)H^T(t)[H(t)P(t^-)H^T(t) + R]^{-1}, \\ \hat{\mathbf{X}}(t^+) &= \hat{\mathbf{X}}(t^-) + K(t)(\tilde{Z} - \hat{Z}), \\ P(t^+) &= [I - K(t)H(t)]P(t^-)[I - K(t)H(t)]^T + K(t)RK^T(t), \end{aligned} \tag{3.17}$$

where H is the measurement geometry matrix, and R is the covariance matrix of the measurement noise. In the particular case of this research, the measurement geometry matrix is defined in (3.10).

Derivation of the Kalman update equations (3.17) was based on both the state vector and measurement probability density functions being Gaussian. The dynamics were also assumed linear, so that if the initial distribution is Gaussian, it will remain Gaussian as it propagates forward-in-time. Below it will be shown that the Step-Back Kalman filter, derived in this work, will match the Kalman filter definition when both the dynamics and measurement models are linear.

A block diagram shown in 3.2 depicts the functionality of the Extended Kalman filter.

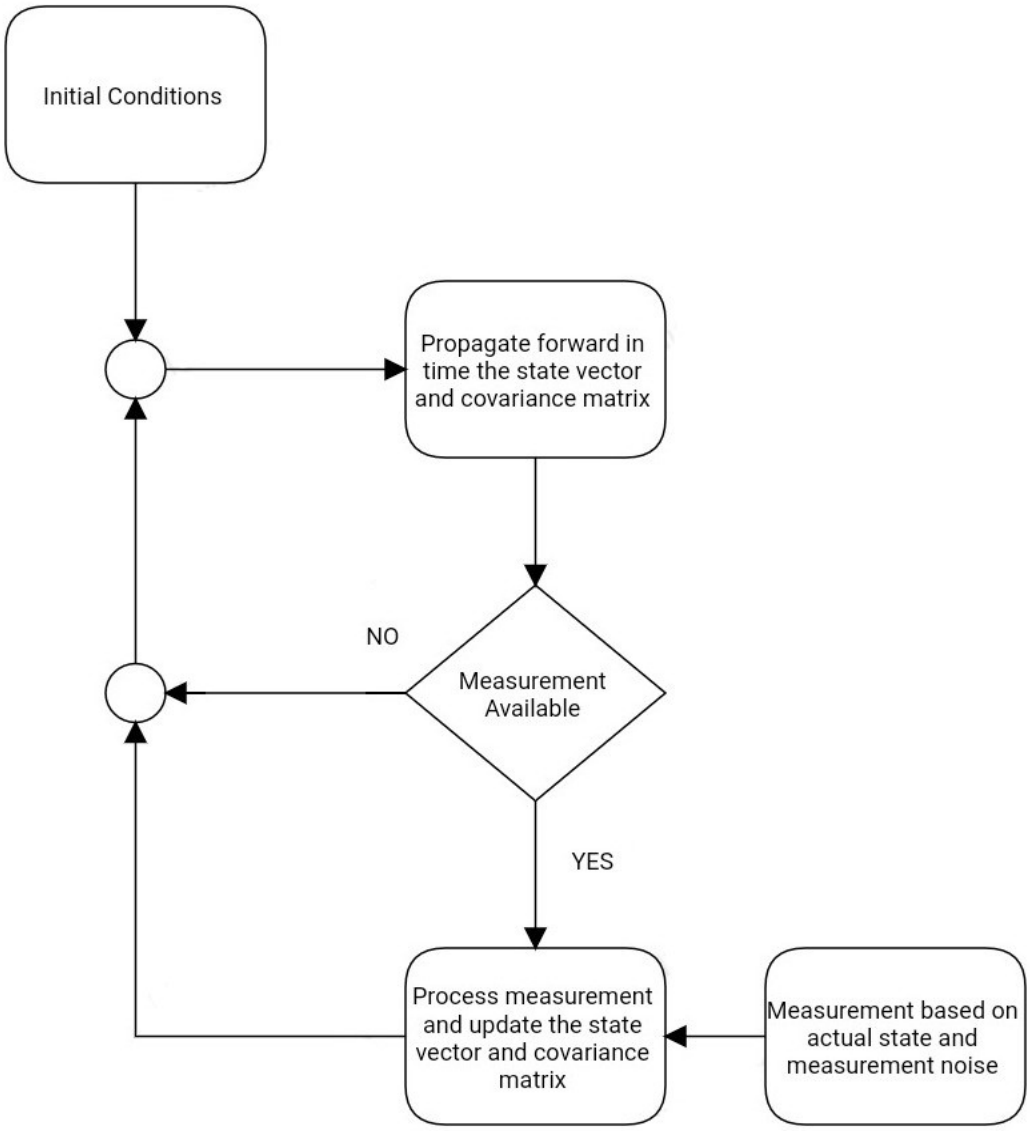


Fig. 3.2: Extended Kalman filter Block Diagram

Block diagrams were created using software by Flowdia Diagrams[1]

[1]Flowdia Diagrams Lite - free download available on Google Play

3.6.2 Particle Filters

A particle filter is essentially a filter in which multiple realizations of the state are generated at the initial start time of the simulation and propagated through the nonlinear orbital dynamics until a measurement is received at which point various strategies of updating can be used depending on the particular particle filter. Some generate particles by random sampling based on the initial condition statistics via,

$$\boldsymbol{\chi}_i = \mathcal{N}\left(\hat{\mathbf{X}}(t_0), P(t_0)\right). \tag{3.18}$$

Others algorithms also use the statistics given by the initial conditions but generate the particles deterministically. For instance the Unscented Kalman filter acquires samples based on the square-root of the covariance matrix and the Gaussian Sum filter acquires nodes based on Gauss-Hermite quadrature. The resultant filter improvements are tallied and run through their own sample statistical analysis in which the sample mean and variance are gathered and represented as heuristics for validity of said algorithm. However, there is an inherent disadvantage to using a particle filter, which is that almost inevitably the algorithm will require a large sum of particles in order to function properly.

3.6.3 Hybrid Kalman-Particle Filter

A hybrid particle Kalman-type filter, which uses simulated trajectories to approximate the covariance matrix more accurately, was also developed in this research effort. In this filtering algorithm the Kalman update is performed using a sample covariance matrix rather than the linearly propagated value carried through in the EKF. This sample covariance matrix yields a more accurate Kalman Gain which resolves the divergence issue to some degree when the measurements are taken at longer time intervals up to a limit at which

point this too also breaks down. Sample statistics are calculated as follows,

$$\begin{aligned} \hat{\boldsymbol{X}}(t_0) &= \frac{\sum_{i=1}^{N} \boldsymbol{\chi}_i}{N}, \\ P(t_0) &= \frac{\sum_{i=1}^{N} \left(\boldsymbol{\chi}_i - \hat{\boldsymbol{X}}(t_0)\right)\left(\boldsymbol{\chi}_i - \hat{\boldsymbol{X}}(t_0)\right)^T}{N-1}. \end{aligned} \tag{3.19}$$

The propagation of the mean value and covariance of the state vector follows exactly as the EKF via (3.12) and (3.15) however, the sample statistics calculated by (3.19) is propagated by the sample covariance calculated from sample particles. The particles are samples of realization of the state based on the statistics at the last point in time when a measurement update was performed. The particles are sampled based on the following,

The number of particles ranges from $i = 1$ to $i = N$ where N is a prescribed number that ideally is larger enough to yield accurate statistical results and yet is not so large as to mandate excessive computation efforts. Each particle is independently generated and then propagated forward in time through dynamical evolution. When a measurement is available to use for an update, the particles are gathered together and used to calculate sample statistics. The block diagram for the HKF is seen in Figure 3.3.

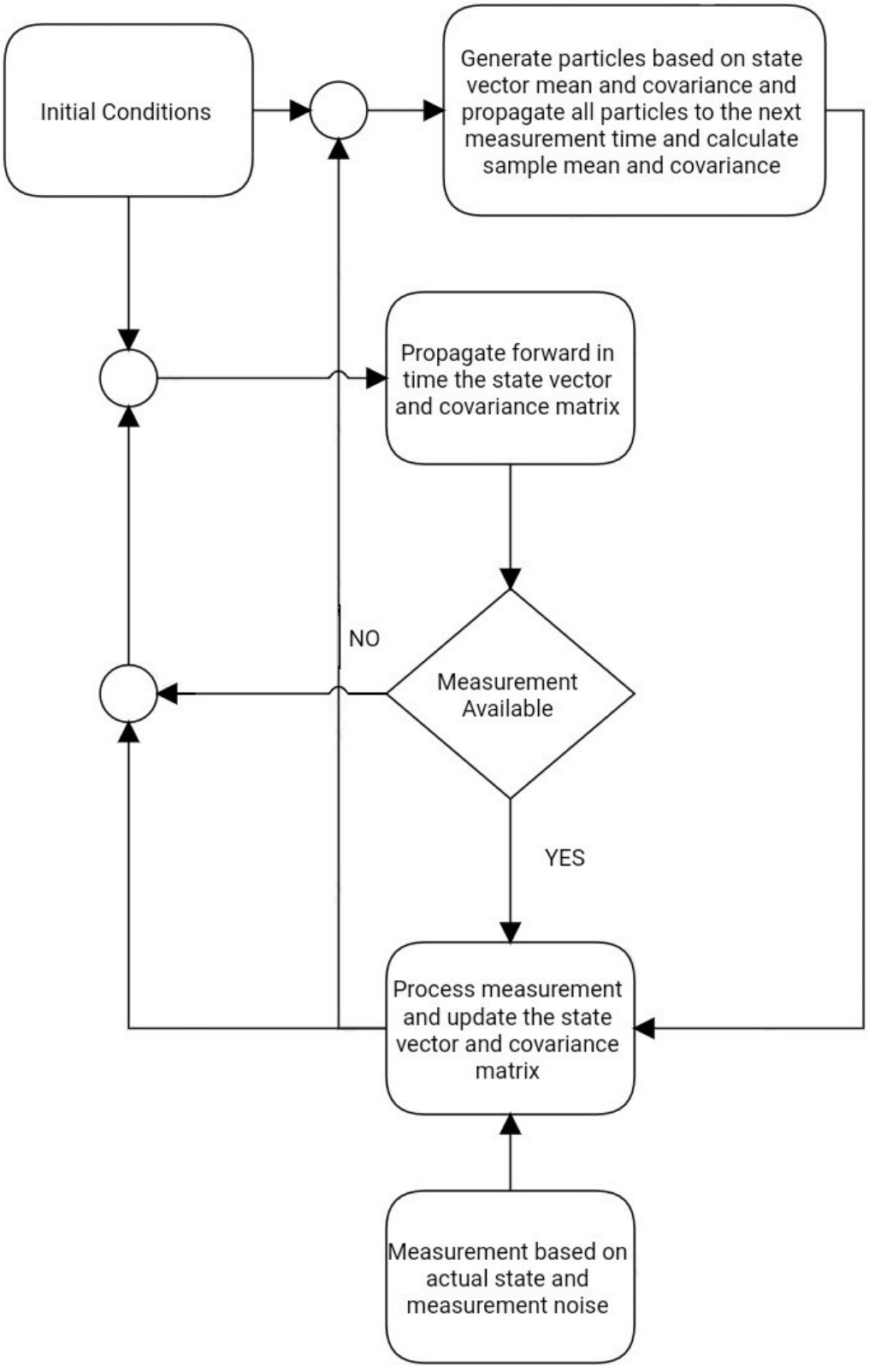


Fig. 3.3: Hybrid Kalman filter Block Diagram

3.6.4 Unscented Kalman Filter

According to [4], the Unscented Kalman filter (UKF) is usually the first option to handle nonlinear effects which cause the EKF to inadequately handle filtering. The UKF is part of a larger class of filtering algorithms called particle filters. The premise behind particle filtering is that realizations of the state space of the RSO are generated, usually randomly based on the probability distribution defined by initial conditions, and propagated through the dynamical equations of motion as well as the measurement model such that sample statistics may be obtained to more closely yield the actual statistics. The UKF essentially uses the particle filter premise with the unique aspect of selecting the particles deterministically based on the mean and covariance defined at the initial time. The particles are essentially chosen along vectors specified by the standard deviation matrix. Using this specific formulation, the mean and covariance can be calculated to third order accuracy for any nonlinear dynamical model [4]. The particles have assigned weight values that have user defined parameters. There are different weights for the first particle with regard to the mean and covariance calculations. The number of particles is always $2L+1$ where L is the size of the state vector. Weight parameters that are user defined are α, β, and κ. These parameters in turn determine parameter λ via,

$$\lambda = \alpha^2 \left(L + \kappa\right) - L. \tag{3.20}$$

From these parameters, the weights for calculating the mean value are given by,

$$\begin{aligned} W_0^{(m)} &= \frac{\lambda}{L+\lambda}, \\ W_i^{(m)} &= \frac{1}{2\left(L+\lambda\right)} \quad i = 1, \ldots, 2L, \end{aligned} \tag{3.21}$$

Likewise, the weights for calculating the covariance matrix are given by,

$$\begin{aligned} W_0^{(c)} &= \frac{\lambda}{L+\lambda} + \left(1 - \alpha^2 + \beta\right), \\ W_i^{(c)} &= \frac{1}{2\left(L+\lambda\right)} \quad i = 1, \ldots, 2L, \end{aligned} \tag{3.22}$$

The location of the particles are parsed from the statistics by the following,

$$\begin{aligned}
\chi_0 &= \hat{\mathbf{X}}, \\
\chi_i &= \hat{\mathbf{X}} + \left(\sqrt{(L+\lambda)\,P}\right)_i, \quad i = 1, \ldots, L, \\
\chi_i &= \hat{\mathbf{X}} - \left(\sqrt{(L+\lambda)\,P}\right)_{i-L}, \quad i = L+1, \ldots, 2L.
\end{aligned} \tag{3.23}$$

After the particles are parsed, they are individually propagated through the Keplerian dynamics as shown in (3.8). At each time step, the mean and covariance are calculated by,

$$\begin{aligned}
\hat{\mathbf{X}}(t) &= \sum_{i=0}^{2L} W_i^{(m)} \chi_i(t), \\
P(t) &= \sum_{i=0}^{2L} W_i^{(c)} \left(\chi_i(t) - \hat{X}(t)\right) \left(\chi_i(t) - \hat{X}(t)\right)^T.
\end{aligned} \tag{3.24}$$

Additive process noise effects are applied in the Unscented Kalman filter algorithm by adding a process noise covariance matrix, P_q, to $P(t)$ at each time step. This requires that at each time step new sigma points need to be generated following (3.23). The process noise covariance matrix was obtained by using a second order integration approximation to obtain a closed-form solution. The following expression (3.25) is the process noise covariance term,

$$P_q = \begin{bmatrix} Q\frac{\Delta t^3}{3} & Q\frac{\Delta t^2}{2} \\ Q\frac{\Delta t^2}{2} & Q\Delta t \end{bmatrix}. \tag{3.25}$$

When a measurement becomes available to process, each particle evaluates a corresponding measurement estimate using (3.9) given the particle state vector. An overall update of the mean and covariance is then performed. The following steps (3.26) show the mathematical

formulation of the UKF measurement update.

$$
\begin{aligned}
\hat{\mathbf{Z}}(t) &= \sum_{i=0}^{2L} W_i^{(m)} \mathbf{Z}\left(\chi_i(t)\right), \\
P_{ZZ} &= \sum_{i=0}^{2L} W_i^{(c)} \left(\mathbf{Z}\left(\chi_i(t)\right) - \hat{\mathbf{Z}}(t)\right) \left(\mathbf{Z}\left(\chi_i(t)\right) - \hat{\mathbf{Z}}(t)\right)^T, \\
P_{XZ} &= \sum_{i=0}^{2L} W_i^{(c)} \left(\chi_i(t) - \hat{X}(t)\right) \left(\mathbf{Z}\left(\chi_i(t)\right) - \hat{\mathbf{Z}}(t)\right)^T, \\
K(t) &= P_{XZ}\left(P_{ZZ} + R\right)^{-1}, \\
\hat{\mathbf{X}}(t^+) &= \hat{\mathbf{X}}(t^-) + K(t)(\tilde{Z} - \hat{Z}), \\
P(t^+) &= P(t^-) - K(t) P_{ZZ} K^T(t),
\end{aligned}
\tag{3.26}
$$

Figure 3.4 shows the block diagram for the UKF.

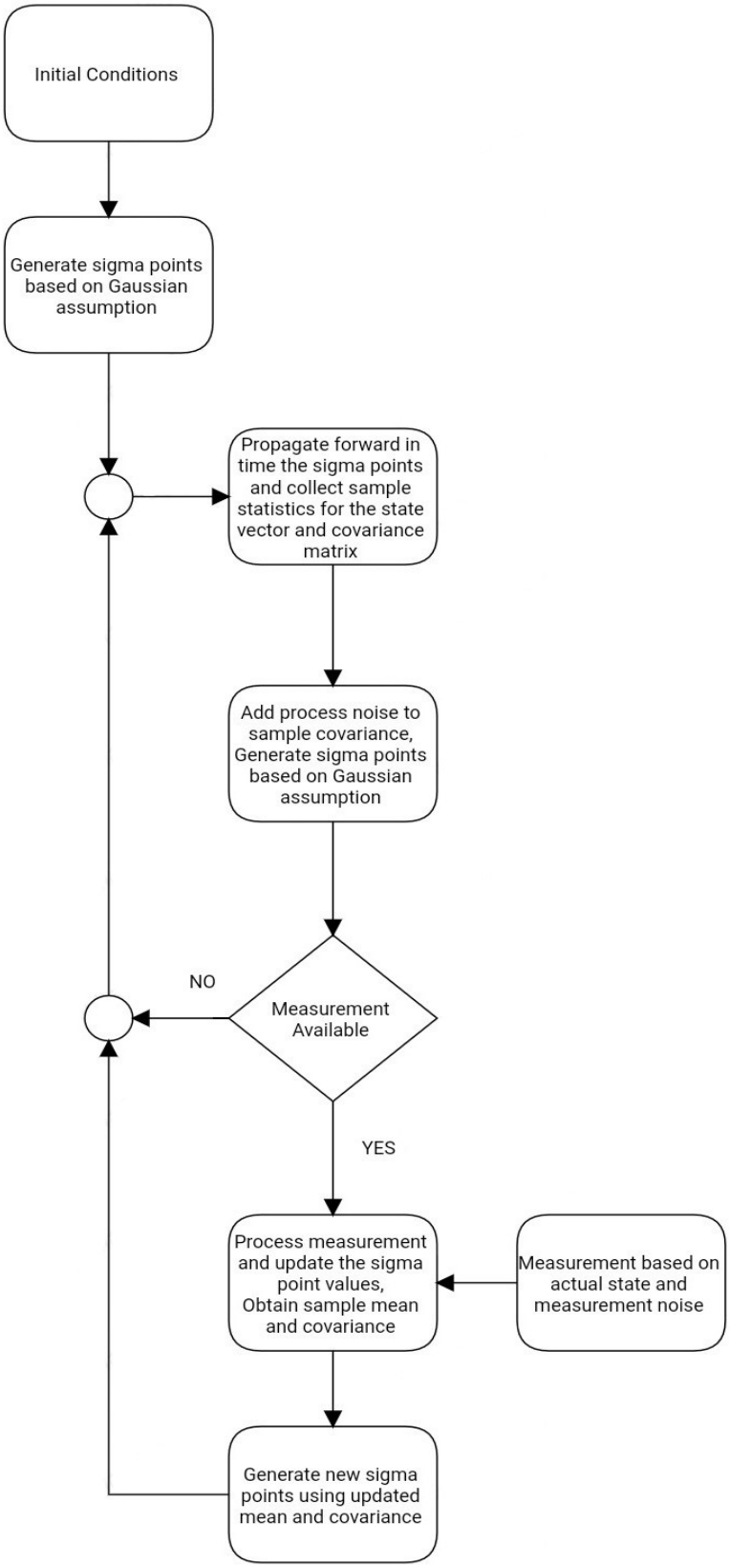


Fig. 3.4: Unscented Kalman filter Block Diagram

3.6.5 Closed Skew-Normal Kalman Filter

Since orbital dynamics are nonlinear, even when initially beginning with a defined Gaussian distribution, as the state vector is propagated through time, the true probability density function accumulates skewness in both the position and velocity vector spaces along the orbital plane. Hence, we are motivated to consider a filtering scheme which can track a distribution of non-zero skewness evolving under Keplerian dynamics. The skewed Kalman filter [23] is a recursive filter for tracking statistics through observation, which replaces the Gaussian approximation of the Kalman filter with a Closed Skew-Normal (CSN) approximation. The CSN distribution generalizes the normal distribution to a larger class which permits the consideration of non-zero skewness. The probability density function $p(\mathbf{X})$, for $\mathbf{X} \in \mathbb{R}^n$, of a CSN distribution with parameters $(\boldsymbol{\psi}, \Omega, \Gamma, \boldsymbol{\nu}, \Delta)$ is given by

$$p(\mathbf{X}) = \frac{\phi(\mathbf{X}; \boldsymbol{\psi}, \Omega)\Phi(\Gamma(\mathbf{X} - \Omega); \boldsymbol{\nu}, \Delta)}{\Phi(\mathbf{0}; \boldsymbol{\nu}, \Delta + \Gamma\Omega\Gamma^T)}, \tag{3.27}$$

where

$$\phi(\mathbf{X}; \boldsymbol{\psi}, \Omega) = \frac{\exp\left(-\frac{1}{2}(\mathbf{X} - \boldsymbol{\psi})^T\Omega^{-1}(\mathbf{X} - \boldsymbol{\psi})\right)}{\sqrt{\left|2\pi\Omega\right|}},$$

is the PDF of a multivariate normal distribution, and

$$\Phi(\Gamma(\mathbf{X} - \boldsymbol{\psi}); \boldsymbol{\nu}, \Delta) = \int_{-\infty}^{\Gamma(\mathbf{X}-\boldsymbol{\psi})} \phi(\mathbf{X}; \boldsymbol{\nu}, \Delta)\, \mathrm{d}\mathbf{X},$$

and

$$\Phi(\mathbf{0}; \boldsymbol{\nu}, \Delta + \Gamma\Omega\Gamma^T) = \int_{-\infty}^{\mathbf{0}} \phi(\mathbf{X}; \boldsymbol{\nu}, \Delta + \Gamma\Omega\Gamma^T)\, \mathrm{d}\mathbf{X}$$

are evaluations of the cumulative distribution function (CDF) of a multivariate normal. The linear skewed Kalman filter has similar formulations as the original Kalman filter [23] and could be used to create an extended form in the same manner as the EKF. However,

there exists an unscented form of this skewed Kalman filter [30]which is referred to as the Closed Skew-Normal Unscented Kalman filter (CSNUKF). The filter algorithm follows a state vector $\mathbf{X}$ with a closed skew-normal distribution having parameters defined at discrete points in time $CSN(\boldsymbol{\psi}_i, \Omega_i, \Gamma_i, \boldsymbol{\nu}_i, \Delta_i)$ and is propagated forward through time following the same methodology as laid out in the section describing the Unscented Kalman filter. The UKF is based on an approximation that the state PDF is Gaussian but since the CSN PDF is not Gaussian, this is not a good approximation. However, it turns out that the equations of the linear skewed Kalman filter as described in an article by Naveau, et. al [23] can be represented by an augmented state space representation which can be propagated assuming a Gaussian a PDF is this augmented space. The augmented state vector $\mathbf{X}_{aug}$ and covariance matrix P_{aug} can be formulated as the following,

$$\begin{aligned} \mathbf{X}_{aug} &= \begin{bmatrix} \boldsymbol{\psi} \\ \boldsymbol{\nu} \end{bmatrix} \\ P_{aug} &= \begin{bmatrix} \Omega & \Omega\Gamma^T \\ \Gamma\Omega & \Delta + \Gamma\Omega\Gamma^T \end{bmatrix} \end{aligned} \tag{3.28}$$

The augmented state space model in equation set (3.28) can now be propagated forward in time following the UKF algorithm and when a measurement is available, it will likewise follow the UKF measurement steps. At each time step when the performance of the filter needs to be evaluated, the CSN parameters will need to be parsed out of the augmented state space and then the mean and covariance of the filter will need to be evaluated using a numerical sampling or integration scheme. The reason for the need to sample or integrate numerically is that the mean and covariance are only equivalent to the CSN parameters $\boldsymbol{\psi}$ and Ω, respectively, if Γ is set to zero. This unique situation means the skewness is zero and the CSN is identical to a Gaussian. However, the skewness will likely not be zero and therefore none of the CSN parameters are identical to the mean and covariance, hence, to acquire the mean and covariance either integration of random sampling will be required.

3.6.6 Gaussian Mixture Model Filter

The issue of divergence still occurs in the UKF and Hybrid filter as a primary result of the state vector density function inheriting skewness over time, thereby further deviating from a Gaussian distribution. Another filter algorithm that follows the particle filter approach is known as the Gaussian Sum filter or the Gaussian Mixture Model filter (GMM). The approach of the GMM is to match the true state PDF by means of approximating as a set of multiple Gaussian distributions each with specific discrete weights.

At the measurement time before the measurement update is applied, the PDF of the GMM is the following,

$$p(\mathbf{X}) = \sum_{i=1}^{N} \omega_i \cdot \phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right).$$

The PDF of the measurement model will use a linearized model following the same methodology in the EKF algorithm. The linearized model will be different for each particle in the GMM since the evaluation of $\hat{\mathbf{Z}}$ will be unique. Let $h(\mathbf{X})$ be the nonlinear function for the measurement.

$$\hat{\mathbf{Z}}_i = h(\boldsymbol{\mu}_i^-).$$

Likewise, the measurement Jacobian H will be unique for each particle. The measurement PDF for each particle will be defined as the following,

$$p(\mathbf{Z}|\mathbf{X}) = \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right).$$

The updated PDF for the GMM when incorporating the measurement PDF is the following,

$$\begin{aligned} p(\mathbf{X}|\mathbf{Z}) &= \frac{\sum_{i=1}^{N} \omega_i \cdot \phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right)}{\int_{-\infty}^{\infty} \left(\sum_{i=1}^{N} \omega_i \cdot \phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right)\right) \mathrm{d}\mathbf{X}} \\ &= \frac{\sum_{i=1}^{N} \omega_i \cdot \phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right)}{\sum_{i=1}^{N} \int_{-\infty}^{\infty} \left(\omega_i \cdot \phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right)\right) \mathrm{d}\mathbf{X}} \end{aligned} \tag{3.29}$$

The Kalman update equations generate a mean value vector and covariance matrix defining a Gaussian distribution which is identical to the Gaussian distribution calculated by taking the product of the two Gaussian distributions for the state and measurement model and normalizing that product. This premise is used to figure a formula for updating the weighting coefficients in the GMM filter.

$$\phi\left(\mathbf{X}; \boldsymbol{\mu}_i^+, P_i^+\right) = \frac{\phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right)}{\int_{-\infty}^{\infty} \phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right) \mathrm{d}\mathbf{X}} \tag{3.30}$$

Since the integral in the denominator is the normalizing factor, it is a constant value and therefore rearranging (3.30) to solve for this normalizing factor will help to identify a formula for weighting factor update.

$$\begin{aligned} &\int_{-\infty}^{\infty} \phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right) \mathrm{d}\mathbf{X} \\ &= \frac{\phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right)}{\phi\left(\mathbf{X}; \boldsymbol{\mu}_i^+, P_i^+\right)} \end{aligned} \tag{3.31}$$

Again, since the normalizing factor is a constant, the ratio on the right-hand side of equation (3.31) will be a constant, therefore it does not matter what value of $\mathbf{X}$ is used. Hence, the variable $\mathbf{X} = \boldsymbol{\zeta}$ will be used for the ratio in (3.31).

$$\begin{aligned} &\int_{-\infty}^{\infty} \phi\left(\mathbf{X}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\mathbf{X} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right) \mathrm{d}\mathbf{X} \\ &= \frac{\phi\left(\boldsymbol{\zeta}; \boldsymbol{\mu}_i^-, P_i^-\right) \phi\left(\hat{\mathbf{Z}}_i + H_i(\boldsymbol{\zeta} - \boldsymbol{\mu}_i^-); \tilde{\mathbf{Z}}, R\right)}{\phi\left(\boldsymbol{\zeta}; \boldsymbol{\mu}_i^+, P_i^+\right)} \end{aligned}$$

Substituting this expression in (3.31) into (3.30) and rearranging leads to the following expression,

$$\begin{aligned}&\phi\left(\mathbf{X};\boldsymbol{\mu}_i^-,P_i^-\right)\phi\left(\hat{\mathbf{Z}}_i+H_i(\mathbf{X}-\boldsymbol{\mu}_i^-);\tilde{\mathbf{Z}},R\right)\\&=\left(\frac{\phi\left(\boldsymbol{\zeta};\boldsymbol{\mu}_i^-,P_i^-\right)\phi\left(\hat{\mathbf{Z}}_i+H_i(\boldsymbol{\zeta}-\boldsymbol{\mu}_i^-);\tilde{\mathbf{Z}},R\right)}{\phi\left(\boldsymbol{\zeta};\boldsymbol{\mu}_i^+,P_i^+\right)}\right)\cdot\phi\left(\mathbf{X};\boldsymbol{\mu}_i^+,P_i^+\right)\end{aligned}\tag{3.32}$$

Now, substituting (3.32) and (3.31) into (3.29)

$$p(\mathbf{X}|\mathbf{Z})=\frac{\sum_{i=1}^N\omega_i\cdot\phi\left(\frac{\phi\left(\boldsymbol{\zeta};\boldsymbol{\mu}_i^-,P_i^-\right)\phi\left(\hat{\mathbf{Z}}_i+H_i(\boldsymbol{\zeta}-\boldsymbol{\mu}_i^-);\tilde{\mathbf{Z}},R\right)}{\phi\left(\boldsymbol{\zeta};\boldsymbol{\mu}_i^+,P_i^+\right)}\right)\cdot\phi\left(\mathbf{X};\boldsymbol{\mu}_i^+,P_i^+\right)}{\sum_{i=1}^N\omega_i\cdot\phi\left(\frac{\phi\left(\boldsymbol{\zeta};\boldsymbol{\mu}_i^-,P_i^-\right)\phi\left(\hat{\mathbf{Z}}_i+H_i(\boldsymbol{\zeta}-\boldsymbol{\mu}_i^-);\tilde{\mathbf{Z}},R\right)}{\phi\left(\boldsymbol{\zeta};\boldsymbol{\mu}_i^+,P_i^+\right)}\right)}\tag{3.33}$$

Hence, the weighting factors will be updated by the following two steps shown in (3.34) and (3.35),

$$\omega_i^+=\omega_i^-\cdot\left(\frac{\phi\left(\boldsymbol{\zeta};\boldsymbol{\mu}_i^-,P_i^-\right)\phi\left(\hat{\mathbf{Z}}_i+H_i(\boldsymbol{\zeta}-\boldsymbol{\mu}_i^-);\tilde{\mathbf{Z}},R\right)}{\phi\left(\boldsymbol{\zeta};\boldsymbol{\mu}_i^+,P_i^+\right)}\right)\tag{3.34}$$

Then normalize the weighting factors,

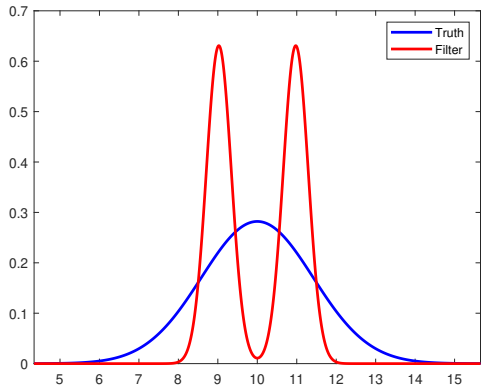
$$\boldsymbol{\omega}=\frac{\boldsymbol{\omega}}{\sum_{i=1}^N\omega_i}\tag{3.35}$$

To give a visual depiction of how the a Gaussian mixture better approximates a non-Gaussian distribution, the following one-dimensional example is used.

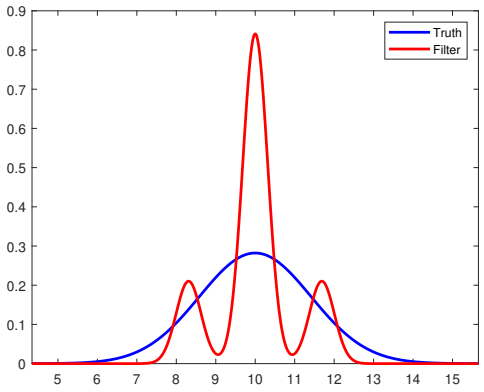
$$\begin{aligned}&\mathbb{E}\left[x\right]=10,\quad\mathbb{E}\left[\left(x-\mathbb{E}\left[x\right]\right)\left(x-\mathbb{E}\left[x\right]\right)^T\right]=2,\quad x=\mathcal{N}\left(10,2\right)\\&y=x-\frac{\sin\left(3x\right)}{5}\end{aligned}$$

The variable x has an associated PDF that is a Gaussian with mean value of 10 and variance 2. When passed into the nonlinear function y, the PDF for the transformation has a shape that could not be closely approximated as a Gaussian. The PDF for y is multi-modal and

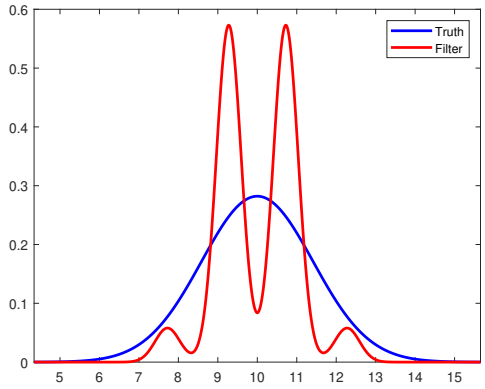
therefore cannot be well-approximated as a Gaussian. Figure 3.5 depicts the matching of the GMM of the Gaussian distribution in the x domain. The PDF in the domain of y, before any measurement is applied in a filtering scheme, the GMM approximation to the true PDF is depicted in Figure 3.6. Figure 3.7 is the block diagram for the GMM.



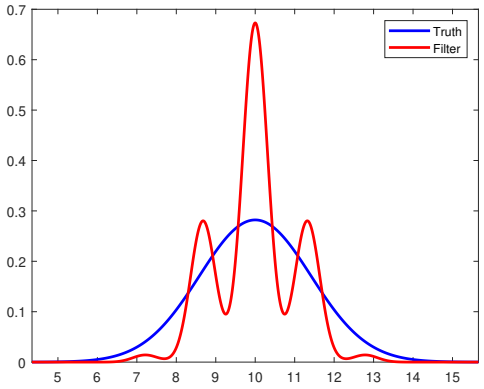
(a) number of nodes = 2



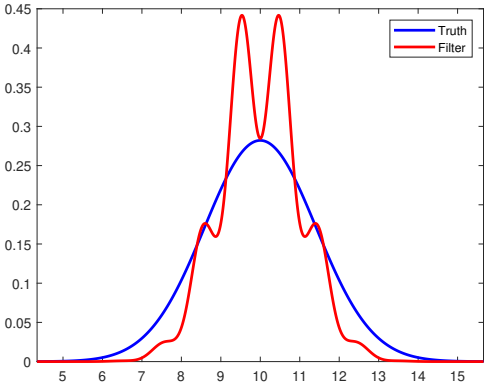
(b) number of nodes = 3



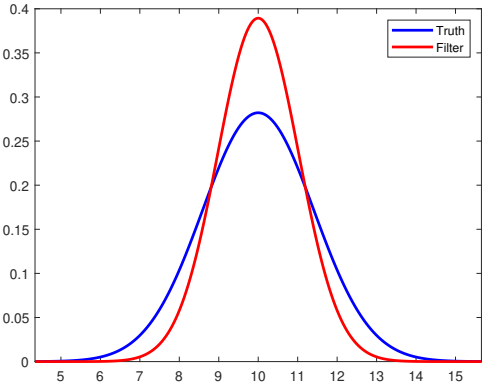
(c) number of nodes = 4



(d) number of nodes = 5

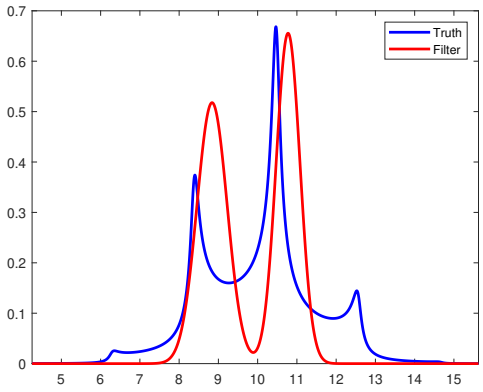


(e) number of nodes = 10

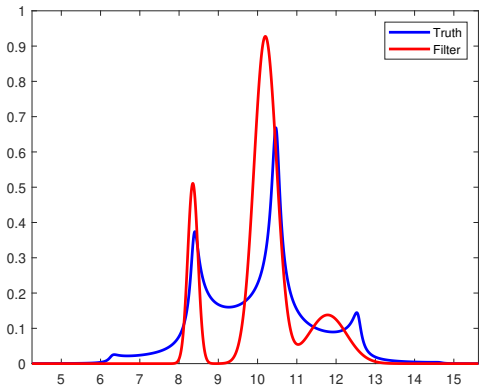


(f) number of nodes = 100

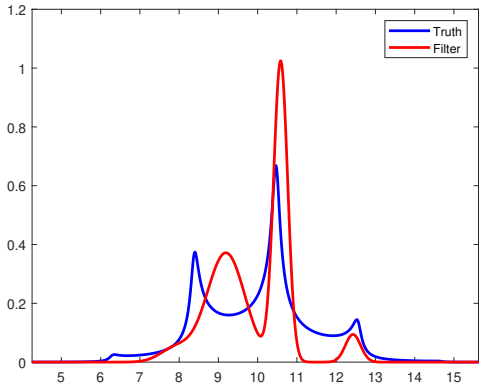
Fig. 3.5: Univariate Gaussian Mixture X-domain



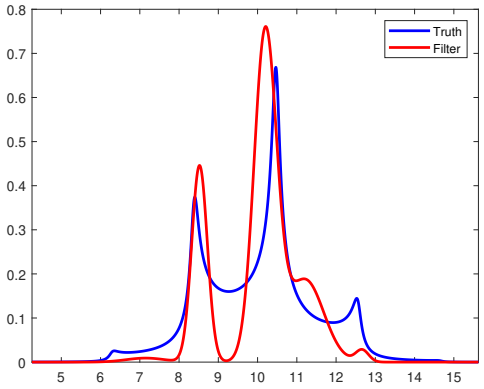
(a) number of nodes = 2



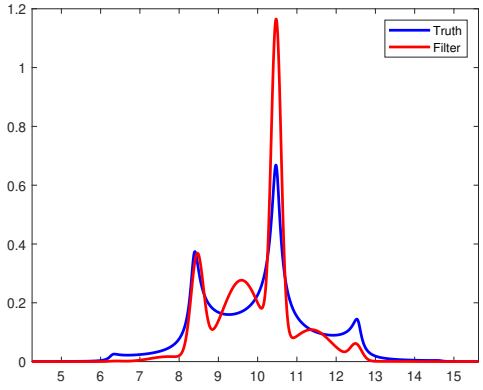
(b) number of nodes = 3



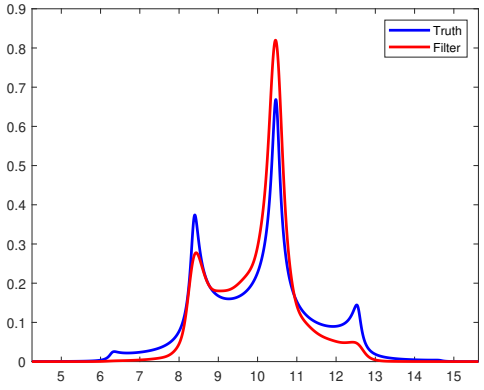
(c) number of nodes = 4



(d) number of nodes = 5



(e) number of nodes = 10



(f) number of nodes = 100

Fig. 3.6: Univariate Gaussian Mixture Y-domain Before Measurement

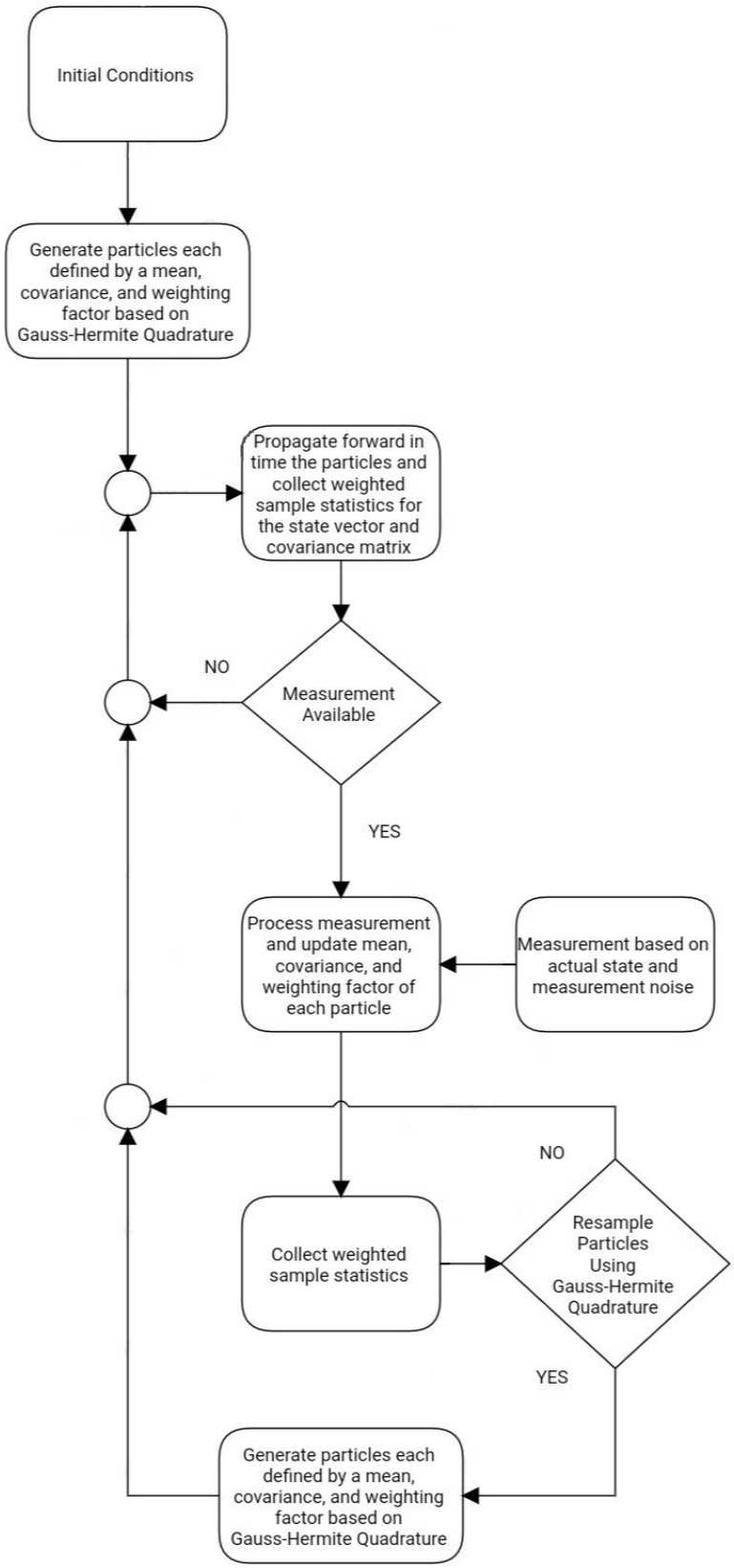


Fig. 3.7: Gaussian Mixture Model filter Block Diagram

3.6.7 Extended Step-Back Kalman Filter

The premise for the Extended Step-Back Kalman filter (ESBKF) is to apply the measurement update at a time in the past when the probability distribution of the state is Gaussian, or at least well-approximated by a Gaussian. The distribution at the start of all the previously described filter algorithms is modeled as a Gaussian. After a measurement update is performed, the resulting mean and covariance are then modeled as describing a new Gaussian distribution. The ESBKF begins exactly the same as the EKF by propagating the state vector estimate and covariance forward in time until a measurement prediction, $\hat{Z}$, is available to process. The measurement Jacobian, $H(t)$ is obtained the same as in (3.10). However, the basis for the step-back formulation is that applying the measurement update to the previous time in the past when the PDF is well-approximated to be Gaussian avoids the error from updating following Kalman update equations (3.17) when a severely skewed PDF is involved. The point in time in the past when the algorithm will step back and apply the measurement update will either be at the initial time of $t = 0$ if the measurement is the first measurement taken, or the point in time will be when the last measurement was received. Since immediately following a measurement update, the PDF is well-approximated as a Gaussian. The chain rule will need to be incorporated in order to map the measurement Jacobian appropriately to the previous point in time. The chain rule maps the measurement Jacobian by the following,

$$\frac{\partial \mathbf{Z}}{\partial \mathbf{X}(t_0)} = \frac{\partial \mathbf{Z}}{\partial \mathbf{X}(t)} \frac{\partial \mathbf{X}(t)}{\partial \mathbf{X}(t_0)} \tag{3.36}$$

Recall that the measurement Jacobian is defined as H (3.10) and the state transition matrix is defined as Φ (3.14). Therefore, following (3.36), the measurement Jacobian with respect to the state vector at time t_0 can be calculated by

$$H(t_0) = H(t)\Phi(t, t_0), \tag{3.37}$$

where H and Φ are given in (3.10) and (3.14), respectively. Also, the contribution of process noise added to the covariance matrix during propagation from t_0 to t is given as,

$$P_q = \int_{t_0}^{t} \Phi(\tau, t_0) G Q G^T \Phi^T(\tau, t_0)\, \mathrm{d}\tau.$$

Applying $H(t_0)$ and P_q to find the Kalman update at time t_0 and corresponding statistics at time t is calculated as seen in (3.38) and (3.39).

$$\begin{aligned} P_{q0} &= \Phi^{-1}(t, t_0) P_q \Phi^{-T}(t, t_0) \\ K(t_0) &= [P(t_0^-) + P_{q0}] H^T(t_0) [H(t_0)[P(t_0^-) + P_{q0}] H^T(t_0) + R]^{-1}, \\ \hat{X}(t_0^+) &= \hat{X}(t_0^-) + K(t_0)(\tilde{Z} - \hat{Z}), \\ P(t_0^+) &= [I - K(t_0)H(t_0)][P(t_0^-) + P_{q0}][I - K(t_0)H(t_0)]^T + K(t_0) R K^T(t_0), \end{aligned} \tag{3.38}$$

For more details on the derivation of these equations, see Appendix B. Finally, we propagate forward to time t and update the ESBKF as follows:

$$\begin{aligned} \hat{\mathbf{X}}(t^+) &= \hat{\mathbf{X}}(t_0^+) + \int_{t_0}^{t} \mathbf{N}\left(\hat{\mathbf{X}}(\tau)\right) \mathrm{d}\tau, \\ \Phi(t, t_0) &= I + \int_{t_0}^{t} F\left(\hat{\mathbf{X}}(\tau)\right) \Phi(\tau, t_0)\, \mathrm{d}\tau, \\ P(t^+) &= \Phi(t, t_0) P(t_0^+) \Phi^T(t, t_0). \end{aligned} \tag{3.39}$$

Recall that F is the Jacobian of the dynamical equations of motion with respect to the state as described in (3.13). For a more detailed analysis of how this formulation was derived by matching exactly to the Kalman filter when both the dynamics and measurement model are linear, see Appendix B. The block diagram for the ESBKF is seen Figure 3.8.

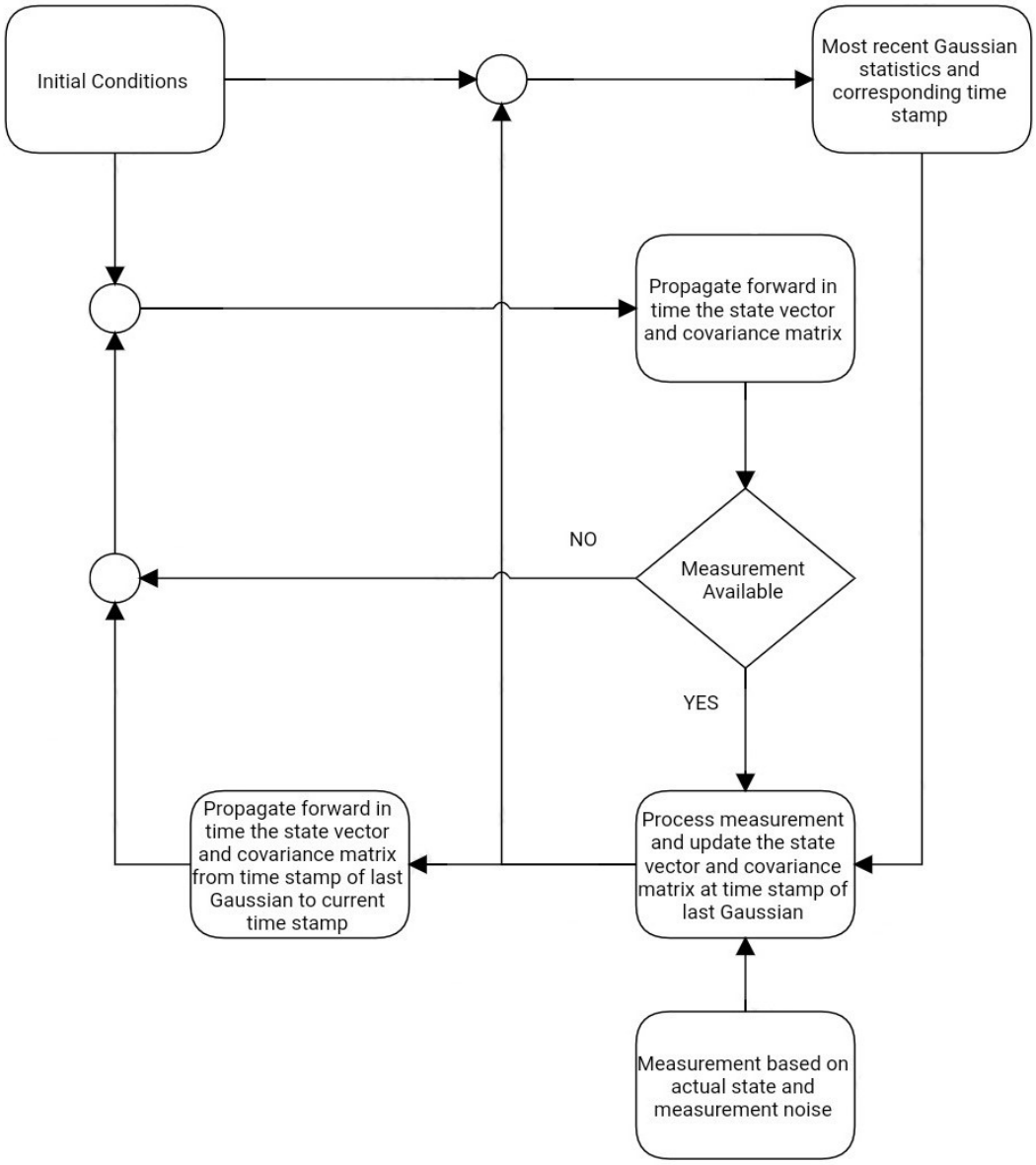


Fig. 3.8: Extended Steb-Back Kalman filter Block Diagram

3.7 Method of Evaluation

Monte Carlo error analysis evaluation is the primary means for determining if a particular filter algorithm can adequately track an RSO. The premise of a Monte Carlo method for evaluation is to acquire statistics by brute-force via random sampling and running a simulation for multiple runs. Typically, the number of runs selected to acquire accurate sample statistics is on the order of several hundred.

Random sampling involves random number generation within each Monte Carlo run and it occurs in three different areas which are the initial true state generation, process noise in dynamical evolution, and measurement noise. At the start of each run, a realization of the true state is generated by sampling from a Gaussian distribution defined by the initial reference state plus a random component based upon the covariance of the initial true state. A filter state is propagated alongside the true state for the RSO as well as the observer satellite through incremental time steps in the simulation. The orbital dynamical model has within it process noise which requires random sampling from a Gaussian distribution based on specified process noise statistics. At specific points in time, angle measurements defining a line of sight from the observer satellite to the RSO are modeled as having measurement noise numerically represented from random sampling from a Gaussian distribution whose mean and covariance are defined by the measuring instrument's accuracy capability.

Each Monte Carlo run will yield unique results of error between the true state and the filter state attempting to track the true state following whichever filtering algorithm is being used. This error between the true state and filter state is referred to as the dispersion. Along with the dispersion, the filter covariance is uniquely obtained in each Monte Carlo run.

After all of the simulation runs are performed and data gathered, the primary heuristic for whether or not a particular filter has performed adequately is to compare the σ values between those obtained by the sample statistics from the dispersion results versus those obtained by taking the average from the filter covariance results. This heuristic is analyzed in two approaches. The first approach is a graphical visualization of 3σ values for each

position and velocity component as a function of time. The second approach is plotting the overlapping index as a function of time, which is the quantity for how closely two Gaussian distributions overlap when taking an L_1-norm calculation of two zero-mean Gaussian PDFs with different σ values.

Visual plots yield insight to the scale of these differences in 3σ values whereas the overlapping index, η, yields a numerical evaluation of filter fidelity. The overlapping index has a threshold in which the filter algorithm is said to become divergent. If η at any point in time falls below 64%, then the filtering algorithm in the simulation is said to be divergent.

This threshold is based on an L_1-norm evaluation of the overlap between two zero-mean univariate normal distributions referred to as an overlapping index [31], [32], [33], [34], [35], [28] having a value between 64% - 100%. The lower limit of 64% was selected from literature in which an overlapping index of 64% in an example was stated as demonstrating two probability distributions as being sufficiently similar [28].

In the RESULTS chapter of this document, plots of η and 3σ are both presented. Since the 3σ plots take up a large quantity of space and for readability reasons, only the Position Downrange is depicted and the remaining other five LVLH components have their 3σ plots presented in APPENDIX F. Analysis will only be presented in the RESULTS chapter and the 3σ plots in APPENDIX F will have limited written context, since they are supplemental plots to the corresponding plots in the RESULTS chapter.

Another metric for filter performance are scatter plots for the RSO position as viewed in the XY plane at the points in time when measurements are received. The simulation has the RSO nominally placed along the equatorial plane which is in this case identical to the XY plane. Depicted in the scatter plots are points of location for the true state, the filter state prior to measurement update, and the filter state after a measurement update has been performed. These scatter plots give visual insight to geometric shapes of the probability distributions.

The rationale for wanting a shape of a probability distribution is because of the foundation of how the Kalman filter was developed. If also the measurement model is linear and

has a Gaussian distribution, the product of these two Gaussian distributions yields the optimal distribution for the state having incorporated all pieces of information [10]. This new distribution was combined from two Gaussian distributions and can be shown to result in a new Gaussian distribution for the state. When dynamics and/or the measurement model are nonlinear, the shape of the distribution will begin to incur an increase in skewness over time and progressively diverge from being well-approximated as a Gaussian.

CHAPTER 4
RESULTS

4.1 Overview

In this chapter, the data collected from Monte Carlo simulations outlined in the last chapter will be presented. The analysis is broken up into four different sections, each giving insight to the performance of the filters. The sections are,

Section 4.2 presents a comprehensive explanation of the three plot types employed in this chapter, which provide different perspectives for the outcomes of each of the filters.

Section 4.3 presents results which demonstrate the cause of why several of the filter algorithms diverge based on results of the PDFs at a specific measurement time. Visual analysis of the PDF at a point when a measurement is processed yields insight in the underlying functionality of these filter algorithms.

Section 4.4 presents results of simulations of the time evolution for all of the filters using two performance metrics: a visual metric in the form of 3σ error/dispersion plots for the

position downrange as a function of time and a numerical metric in the form of overlapping indices for all of the LVLH components as a function of time. This section is broken up further into subsections separated by the measurement times used in the simulation.

Section 4.5 presents the computation times for each of the filters. The computation times are presented as a ratio relative to the time required for the EKF to compute a simulation. Computation times are presented in this manner because the goal is to investigate how well each filter performs in comparison with EKF, not only with regard to accuracy but with speed as well.

Section 4.6 describes in detail why the CSNUKF was pursued as the original focus of this research and the reasons it was not implemented into the simulation process.

Lastly, Section 4.7 presents a summary section will detail conclusions derived from analyses in the aforementioned sections in this chapter.

4.2 Details of Plots

There are three plot types presented in this chapter. These plots are enumerated below followed by a description of each.

1. Positive and negative 3σ values obtained by two methods.
2. Overlapping index based on σ values obtained by two methods.
3. Scatter plot of samples obtained from multiple probability distributions at a specific measurement time.

Plots containing 3σ values over time compare two standard deviations obtained through different calculations using Monte Carlo simulations. The first calculation determines the error between the true state and the filter state for each Monte Carlo run, from which sample standard deviations are obtained. The second calculation determines the average of the covariance matrices tracked by the filter, takes the square-root of the major diagonal elements, and obtains the standard deviation.

The multiplication of the standard deviation by a factor of 3 is a widely used heuristic since approximately 99.7% of all results for a normal distribution should lie within 3σ error

bounds, which effectively translates to the range for all results. These plots present visual insight into the performance of each filter by looking at how close the two standard deviation plots match. The overlapping index, η, gives a numerical metric of success or failure for each filter.

Plots of η give a direct numerical evaluation of whether the filter is convergent or divergent based on whether or not η falls below 64% at any time in the simulation. Each plot contains a calculation of η for each of the six elements, position altitude, downrange, and crosstrack and velocity altitude, downrange, and crosstrack.

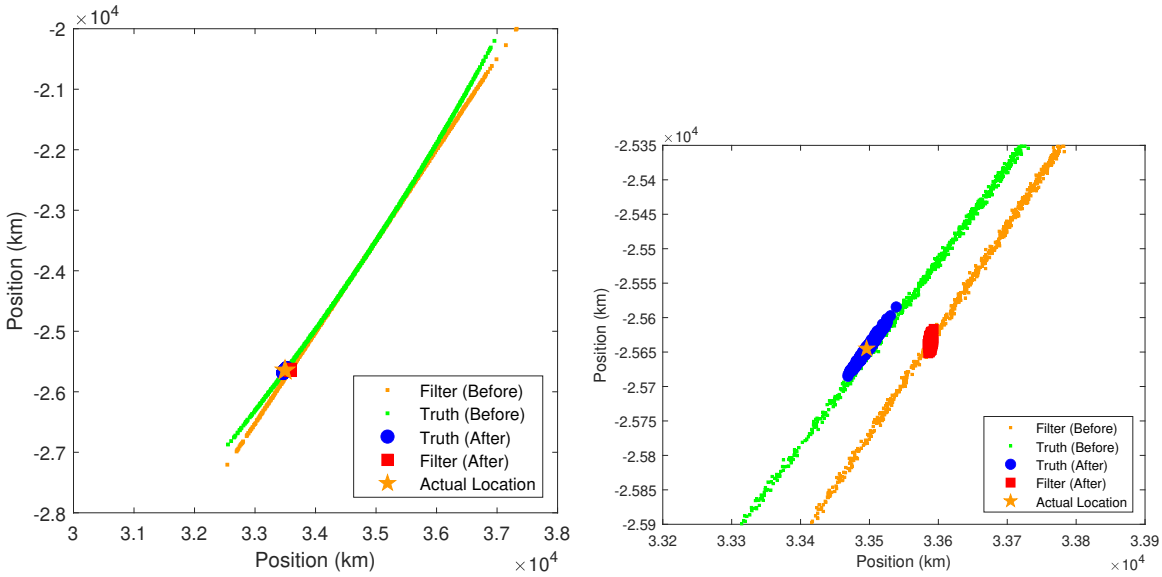
Lastly, scatter plots illustrate the PDFs underlying the filters. These plots provide visual evaluation of what is occurring with regard to the PDFs of the state when a measurement is processed. Essentially, scatter plots show the PDF configuration for a single Monte Carlo run at the time of measurement. This provides context for why the overlapping index yields a failure or success.

This chapter presents 3σ plots that exclusively depict the position downrange component of the RSO. This approach was adopted to enhance the readability of the plots while retaining the key information on the divergence behavior. The position downrange component was found to be a reliable indicator of the divergence behavior, which is why it was selected to be presented in this chapter. APPENDIX F presents 3σ plots for all six LVLH elements and is intended to be used as supplemental information.

4.3 Probability Densities at Measurement Update

Using the EKF for estimating or tracking an RSO in GEO is subject to limitations due to the duration between measurements before it functionally fails. The source of this failure was found to be nonlinear effects present in both the dynamics and the measurement model. The EKF assumes a Gaussian distribution for the state PDF. However, as the state propagates through nonlinear orbital dynamics, this assumption becomes a less accurate approximation for the true PDF. The longer propagation continues without an update from a discrete measurement, the more the true PDF will accrue skewness and will no longer be adequately approximated by a Gaussian PDF. As a result, the EKF will make an inaccurate

update for the RSO, placing it at a higher altitude. This has cascading effects since, at a higher altitude, the filter estimate of the RSO would move at a slower speed, which exacerbates the discrepancy between the filter estimate and the true position and velocity of the RSO.



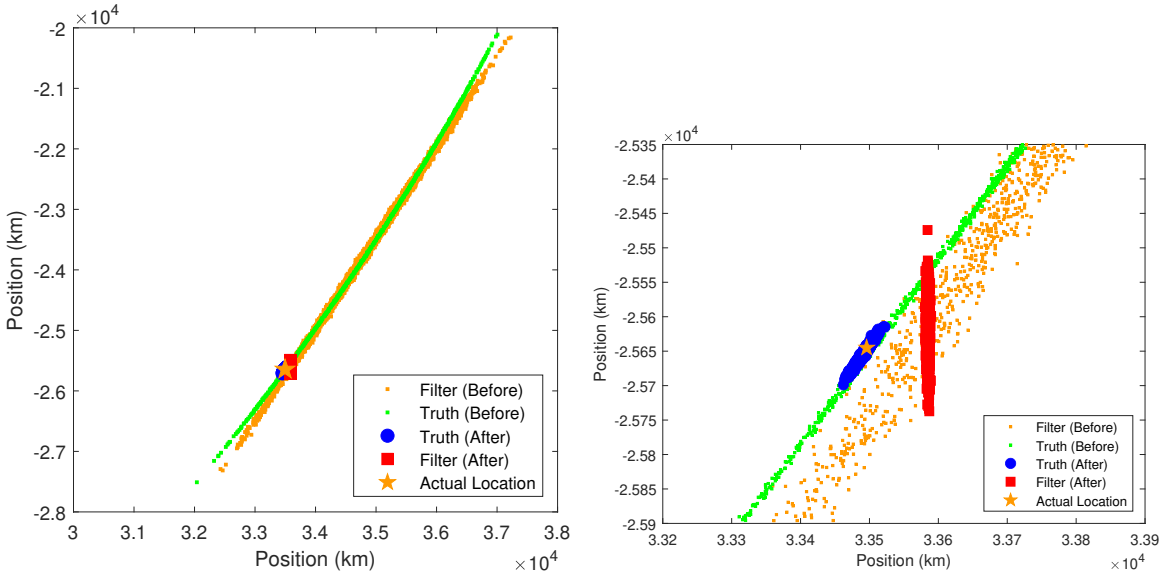
(a) Position XY plane zoomed out view (b) Position XY plane zoomed in view

Fig. 4.1: EKF probability density at 140 hours, OBS 1 Only

The EKF updates the estimates for the mean value and the covariance matrix of the state vector by placing the updated PDF in the incorrect location, as seen in Figure 4.1. The EKF has been using a linearized approximation for the dynamical model and an assumed Gaussian PDF throughout the propagation and therefore when a measurement is processed to update the filter mean and covariance, the EKF positions the updated PDF at the intersection of the measurement model PDF and the Gaussian PDF of the state. This leads to the filter estimate being placed at a higher altitude than if the filter had used the intersection of the actual non-Gaussian PDF of the state with the measurement model PDF, which exhibits skewness.

As shown in Figure 4.1, the PDF of the filter state prior to the measurement update is based on a linearized model and fails to capture the skewed geometry of the actual PDF.

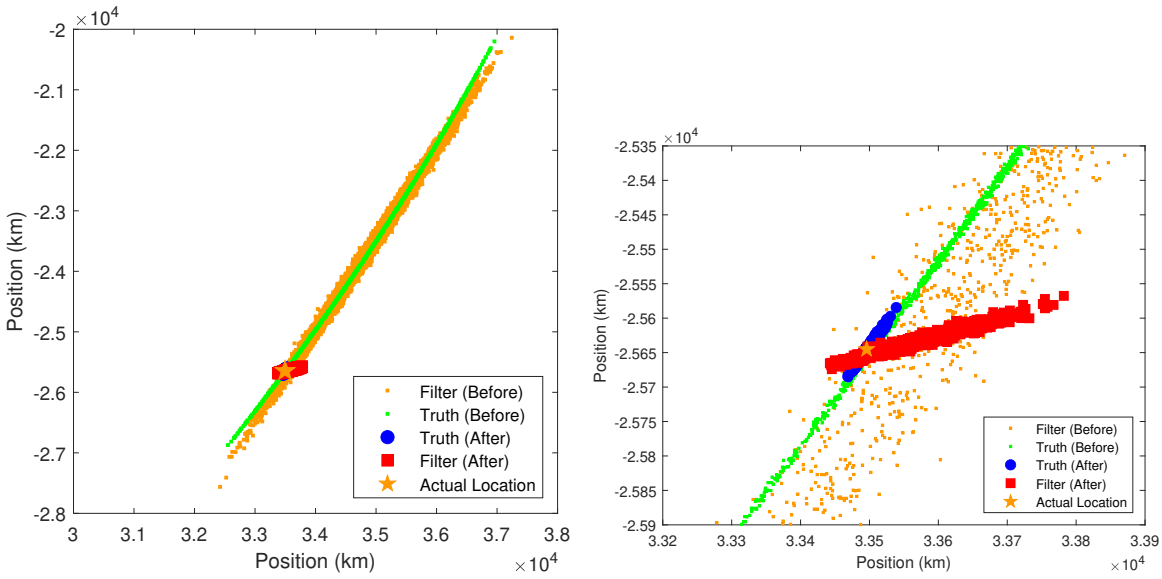
A Gaussian distribution does not accurately represent the tail-ends of the true PDF, and consequently, the intersection with the measurement distribution will place the estimate of the RSO at a higher altitude than it actually is. This higher altitude leads to a lower speed for the RSO in the filter estimate, which exacerbates divergence from reasonable RSO estimation as time progresses.



(a) Position XY plane zoomed out view (b) Position XY plane zoomed in view

Fig. 4.2: HKF probability density at 140 hours, OBS 1 Only

The updated state PDF from the HKF is situated at a higher altitude than the actual location of the RSO, similar to the behavior of the EKF. However, the HKF uses sample statistics, obtained from the true PDF before the measurement, as opposed to the linearly propagated statistics of the EKF. As a result, the covariance of the filter PDF after the measurement update is larger for the HKF compared to the EKF, as illustrated in Figure 4.2. The larger covariance can assist in compensation for errors in the estimate of the filter mean value, since the true location of state is closer to the region of higher probability in the filter PDF.

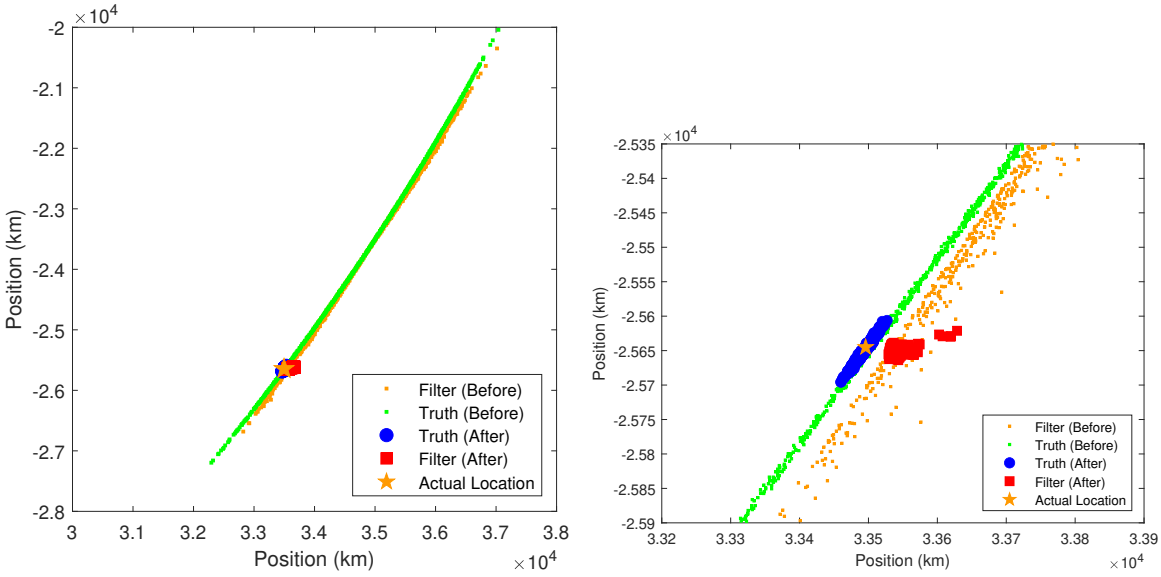


(a) Position XY plane zoomed out view

(b) Position XY plane zoomed in view

Fig. 4.3: UKF probability density at 140 hours, OBS 1 Only

The UKF also situates the filter estimate at a higher altitude. However, similar to the HKF, it also has a larger covariance of the filter PDF after the measurement update, as seen in Figure 4.3. Sigma points in the UKF, having been propagated through the nonlinear dynamics, follows the same heuristic as the HKF of using sample particles to acquire a more accurate mean and covariance when processing the measurement.

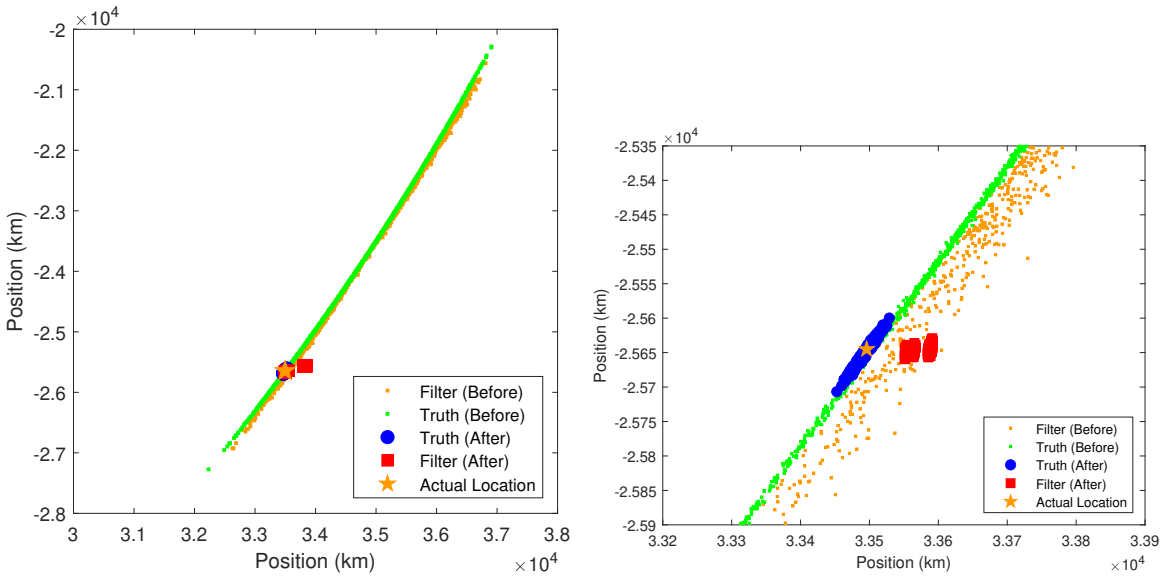


(a) Position XY plane zoomed out view

(b) Position XY plane zoomed in view

Fig. 4.4: GMM (N=2^6) probability density at 140 hours, OBS 1 Only

As depicted in Figure 4.4, the GMM generates a post-measurement estimate of the RSO that is at a higher altitude than the true location but not as significantly as the EKF. These results are with only 2 nodes per dimension, which translates to 64 nodes or particles in total.

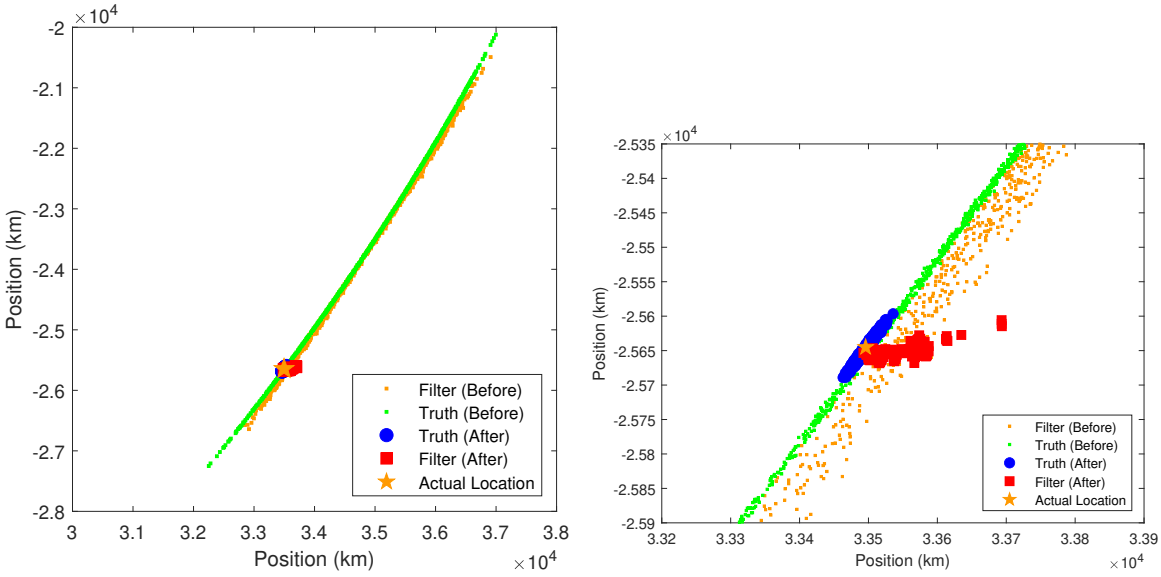


(a) Position XY plane zoomed out view

(b) Position XY plane zoomed in view

Fig. 4.5: GMM (N=5^6) probability density at 140 hours, OBS 1 Only

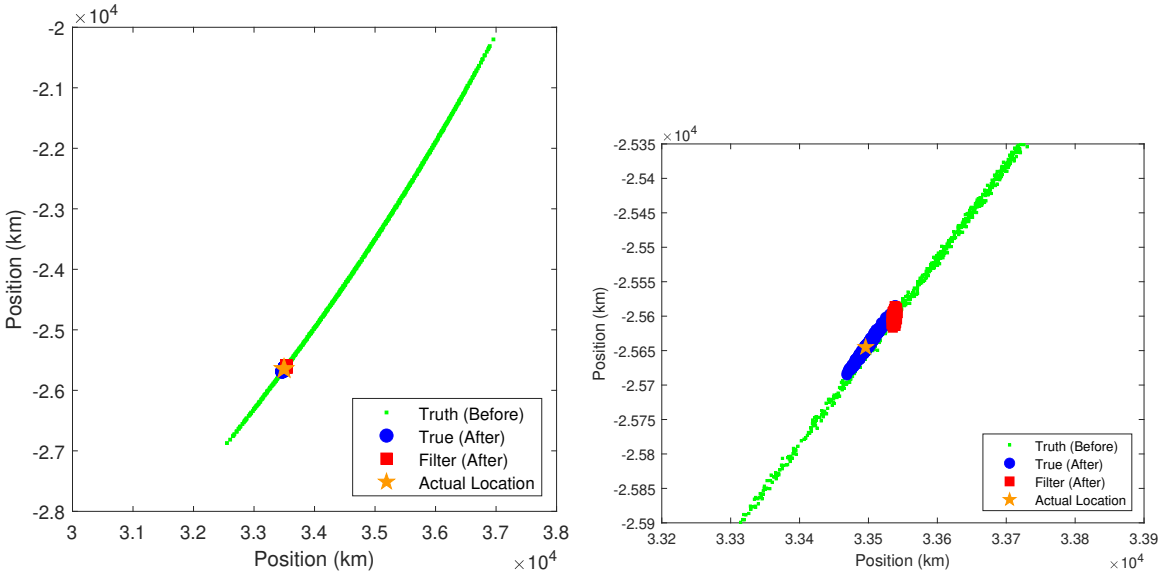
The post-measurement filter PDF in Figure 4.5 was created from the GMM using 5 nodes per dimension and the increase in nodes or particles made it so that a post-measurement estimate of the RSO was closer in altitude to the actual location when compared with Figure 4.4.



(a) Position XY plane zoomed out view (b) Position XY plane zoomed in view

Fig. 4.6: GMM (N=10^6) probability density at 140 hours, OBS 1 Only

The filter PDF after the measurement update, in Figure 4.6, was created from the GMM using 10 nodes per dimension, which has the effect of centering PDF closer to the true location of the RSO than what is seen in Figure 4.5. There can be more confidence the filter estimate is closer to the actual location.



(a) Position XY plane zoomed out view (b) Position XY plane zoomed in view

Fig. 4.7: ESBKF probability density at 140 hours, OBS 1 Only

The ESBKF reduces the effects of nonlinearity with respect to the dynamics, however, there remains nonlinearity in the measurement model. This conclusion is corroborated from the PDFs in Figure 4.7 in which the location of the RSO after the measurement update is offset from the truth and has a smaller covariance. An item of future work will be to develop a technique into the Step-Back Kalman filter methodology that would account for nonlinear measurements. This may be investigating an unscented form of the SBKF, for example.

It should noted that in Figure 4.7, there is not a Filter (Before) PDF shown, which is because the measurement is update at $t = 0$ in the ESBKF. The filter estimate is then propagated to $t = 140$ hours and the resultant PDF for the Filter (After) is shown in Figure 4.7.

4.4 Time Evolution of Standard Deviations and Overlapping Indices

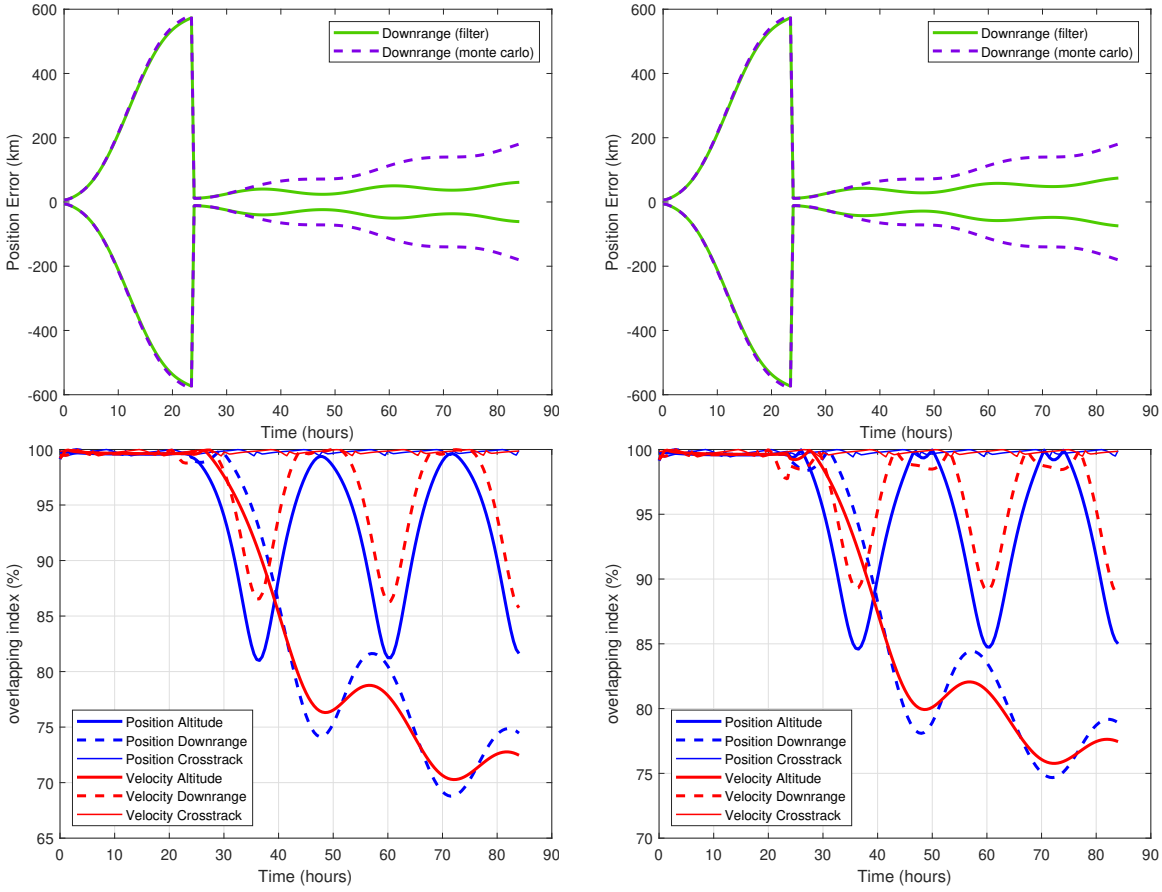
This section presents the simulation results of 3σ and η as functions of time for all five filter algorithms. Subsections separate the results based on the time at which the measurement is processed. There were three measurement times simulated, which were 24, 70, and 140 hours. Within each subsection, results for the four OBS configurations are presented. For each OBS configuration, three plots are shown, which are the following,

1. 3σ and η plots for the EKF and UKF
2. 3σ and η plots for the HKF and GMM
3. 3σ and η plots for the ESBKF

Each plot depicts the 3σ position downrange component of the RSO state, which is placed directly above the η plot to facilitate visual comparison of their time evolution. The η plot displays all six components of the RSO in the LVLH coordinate system. The placement of the plots facilitates in observing significant jumps in both plots together, when the measurement is processed. Additional corresponding 3σ plots for all six LVLH components are presented in APPENDIX F. Relevant links to the plots in APPENDIX F will be provided throughout this section.

4.4.1 Measurement Time of 24 Hours

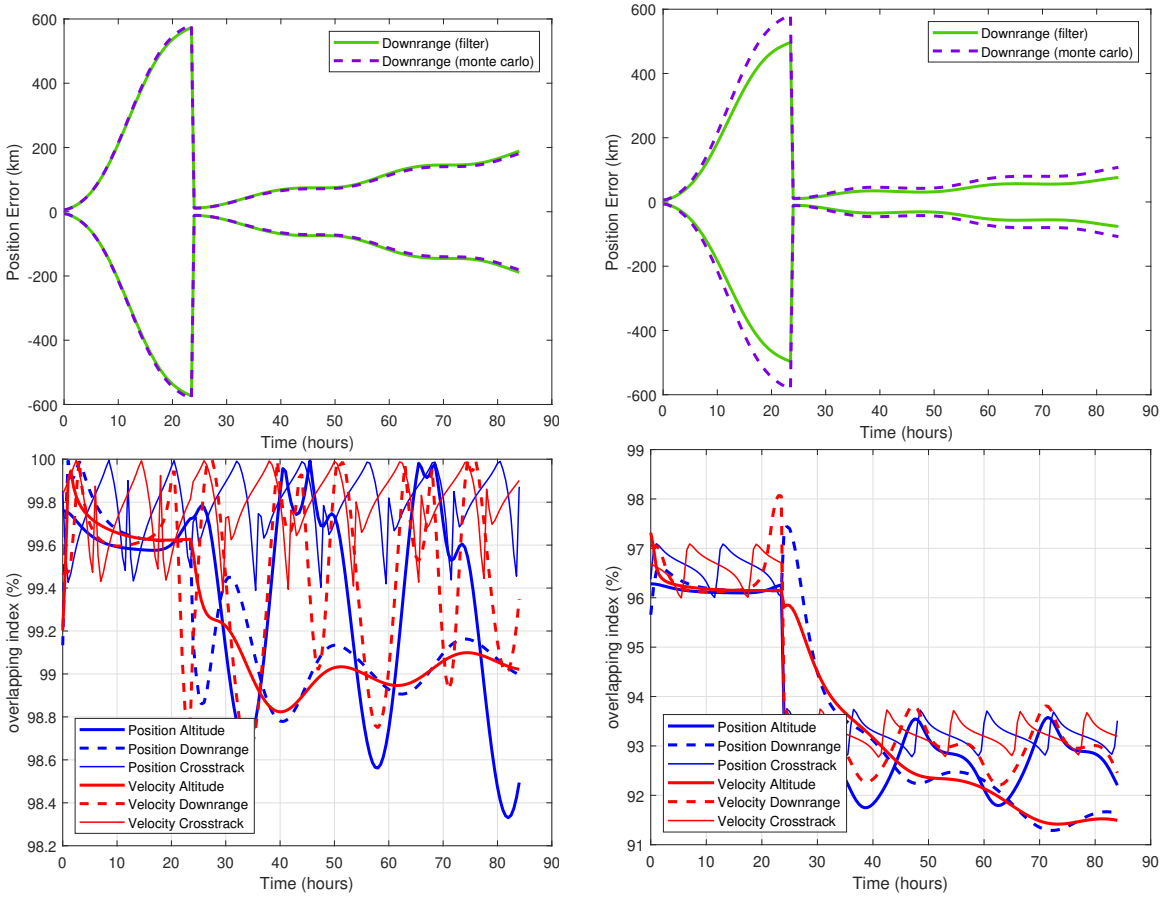
In a scenario where an RSO in GEO is tracked using space-based optical measurements, and the duration between measurements is not too long, both the EKF and the UKF appear to track the RSO accurately. Figure 4.8 shows a progressive deviation of the position downrange and the velocity altitude η lines. It is worth noting that this is also a reason why the position downrange was selected to be presented in the 3σ plots in this chapter.



(a) EKF (b) UKF

Fig. 4.8: 3σ and η EKF and UKF Measurement at 24 Hours, OBS 1 Only

Figure 4.8, illustrates that both the EKF and UKF had overlapping indices that were above the 64% threshold for convergence. When a measurement is processed at 24 hours the EKF and UKF were accurate. See Figures F.1 and F.25 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.9: 3σ and η HKF and GMM Measurement at 24 Hours, OBS 1 Only

The HKF and GMM tracked the RSO accurately when a measurement updated at 24 hours, as shown in Figure 4.9. It was expected that both of these filters would perform well, particularly at a reasonable early measurement time. The EKF suffers from linear approximations of dynamics and assumptions of Gaussian distributions, which is why it loses accuracy quickly. The UKF is forced to reset sigma points at every time step, due to the discrete process noise covariance matrix being added, which assumes a Gaussian distribution from which to obtain sigma points. However, both the HKF and GMM use particles or nodes which do not need to be reset every time step which improves the measurement update accuracy. See Figures F.13 and F.37 for the results of 3σ values for all LVLH components.

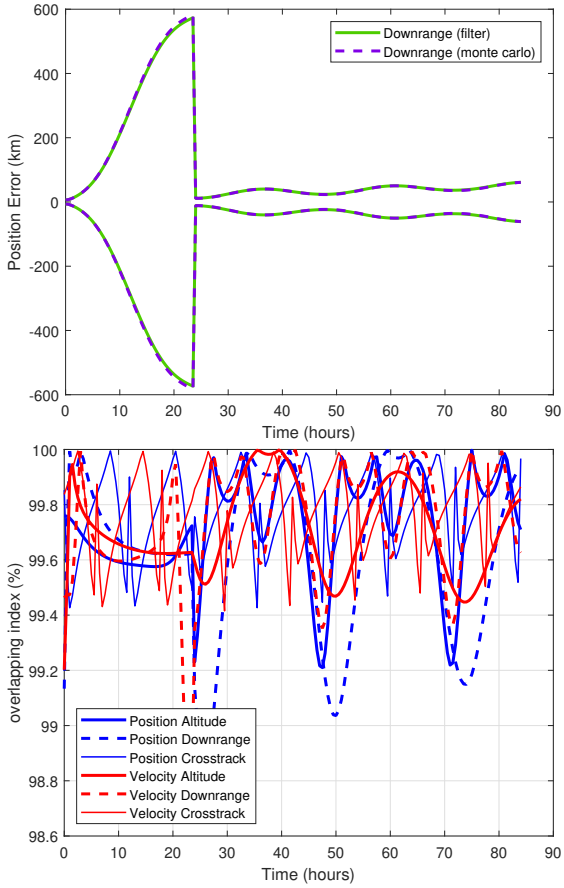
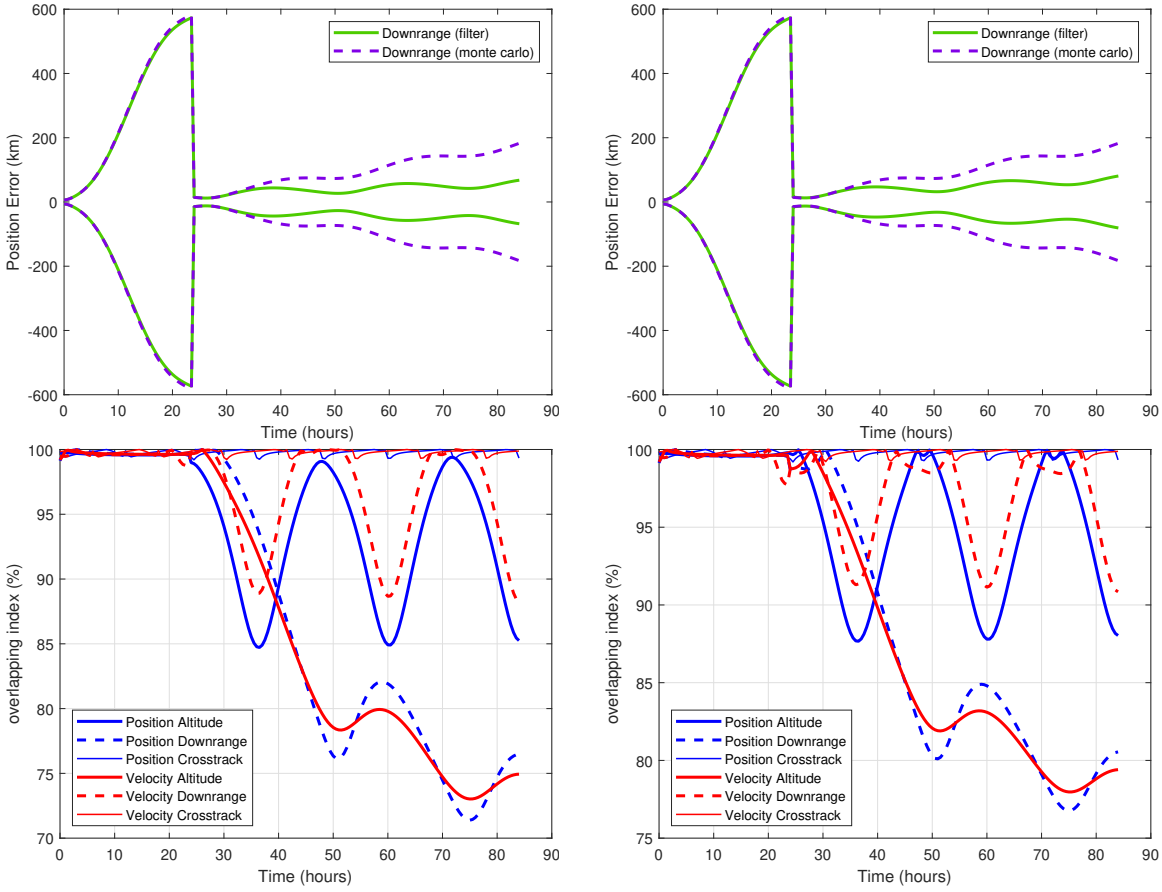


Fig. 4.10: 3σ and η ESBKF Measurement at 24 Hours, OBS 1 Only

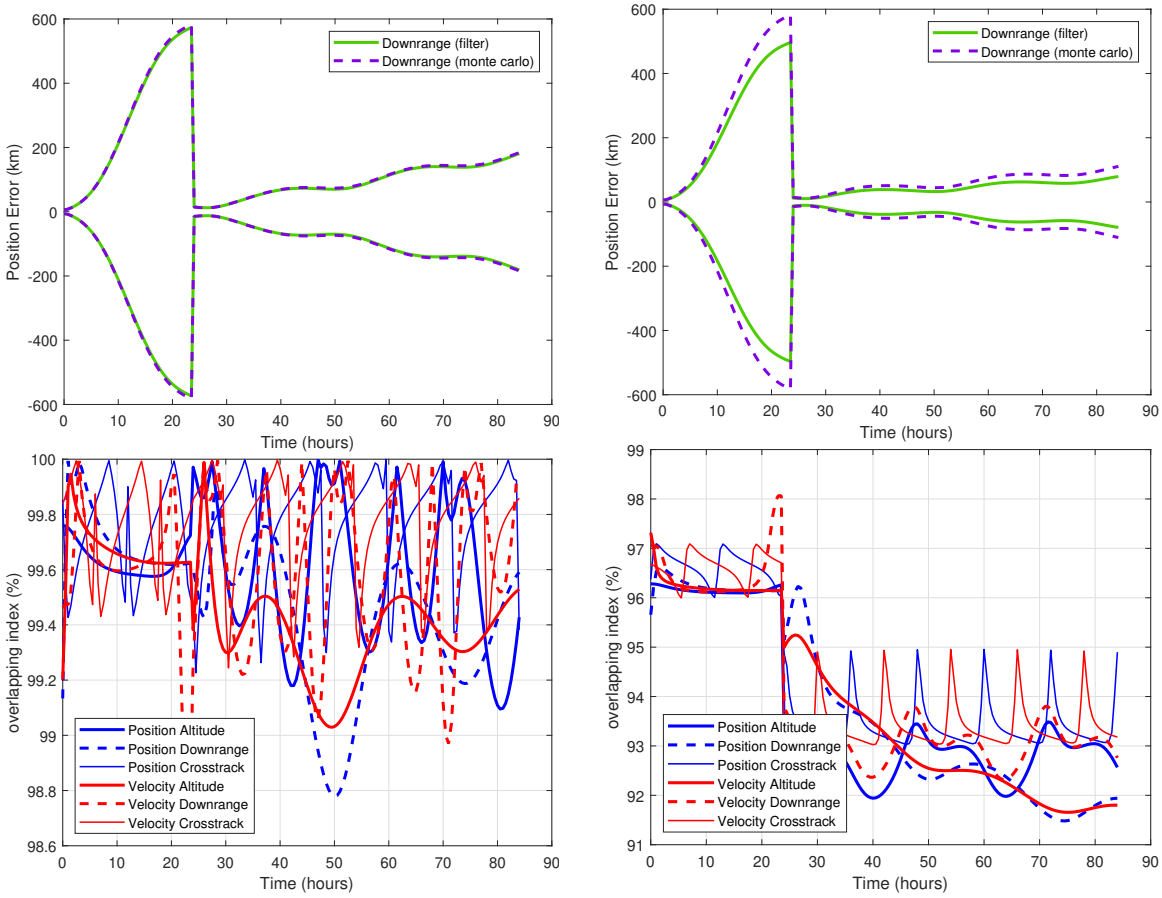
The ESBKF tracked the RSO very accurately. Both the HKF and ESBKF had η values that were nearly 100% but the ESBKF also had a smaller 3σ value after the measurement update, as shown in Figure 4.10. The 3σ values increase faster for the HKF after the measurement because the size of the standard deviation is larger than the ESBKF. Having a PDF defined with a smaller covariance will take longer to accrue skewness. This will also be illustrated in other results in this chapter. See Figure F.49 for the results of 3σ values for all LVLH components.



(a) EKF (b) UKF

Fig. 4.11: 3σ and η EKF and UKF Measurement at 24 Hours, OBS 2 Only

The performance of both the EKF and UKF did not exhibit a significant change when using the OBS configuration with a higher altitude, as indicated in Figure 4.11. It was expected that the OBS 2 configuration would yield more accurate results than OBS 1, since a higher altitude for the OBS places it closer to the RSO. Measurement errors reside in the angle domain, and therefore if the OBS is in a closer proximity to the RSO, then the errors in the position would be smaller. A decrease in position errors should intuitively result in a more accurate filter update for the estimate. This was not observed in Figure 4.11, which implies that the proximity of the RSO is not as important. See Figures F.2 and F.26 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.12: 3σ and η HKF and GMM Measurement at 24 Hours, OBS 2 Only

Figure 4.12 illustrates that the HKF and GMM did not exhibit significant changes when using OBS 2 instead of OBS 1 configurations. Both filters accurately tracked the RSO. These filters, along with the ESBKF, were expected to be the most accurate, and the results for measurements processed at 24 hours support this conclusion. Please refer to Figures F.14 and F.38 for the results of 3σ values for all LVLH components.

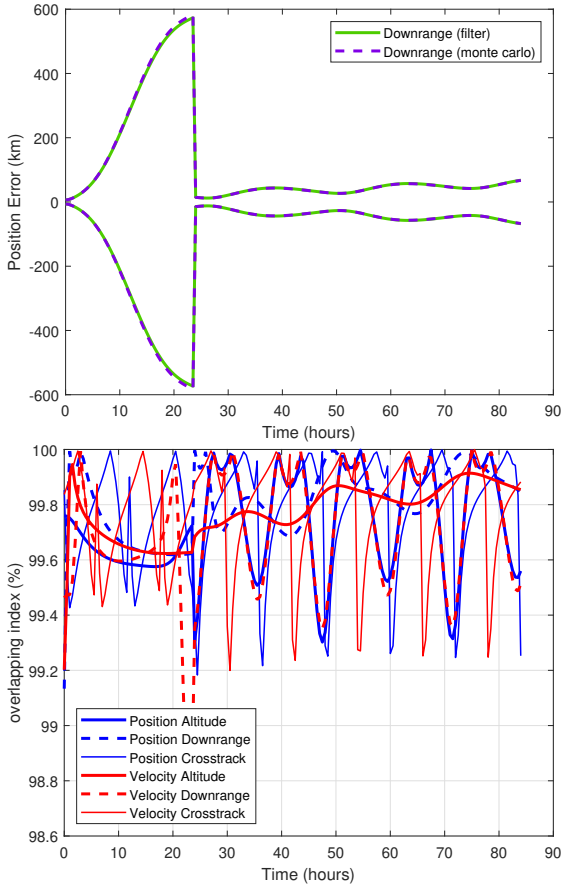
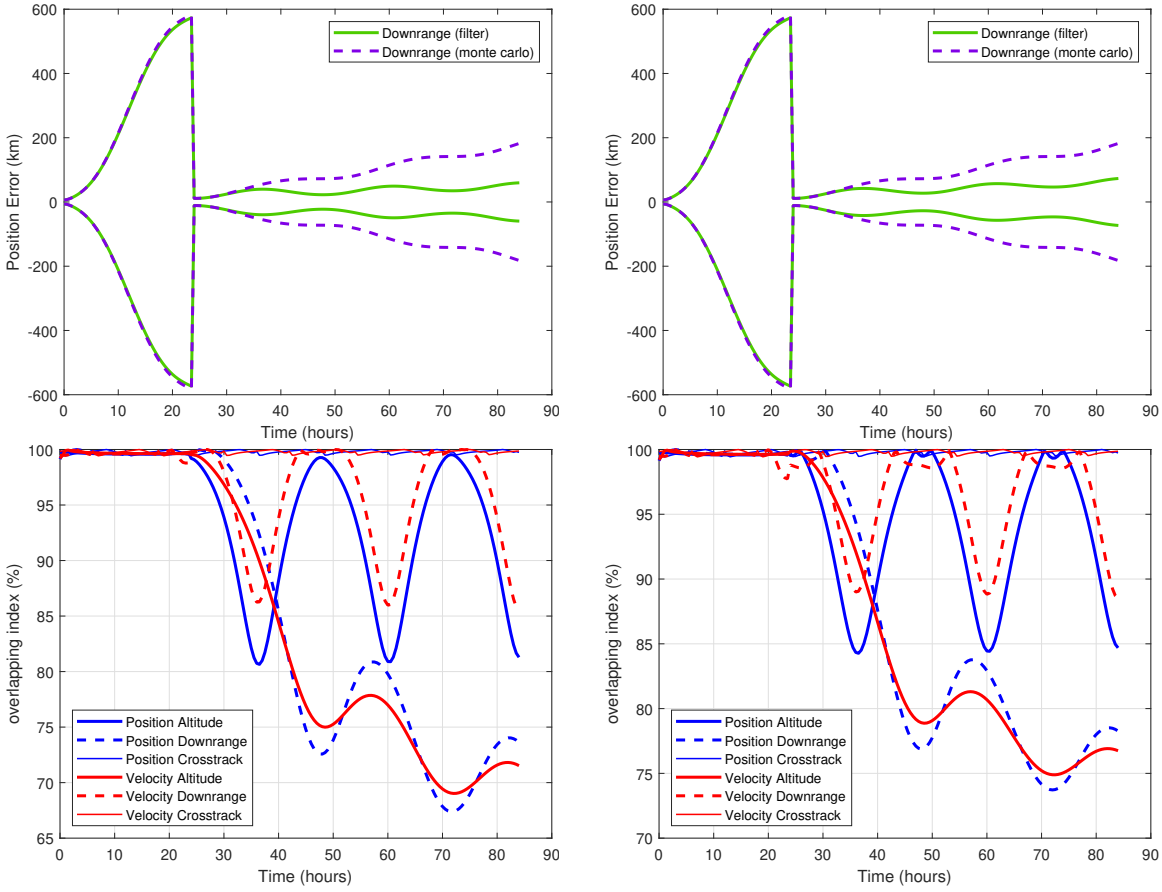


Fig. 4.13: 3σ and η ESBKF Measurement at 24 Hours, OBS 2 Only

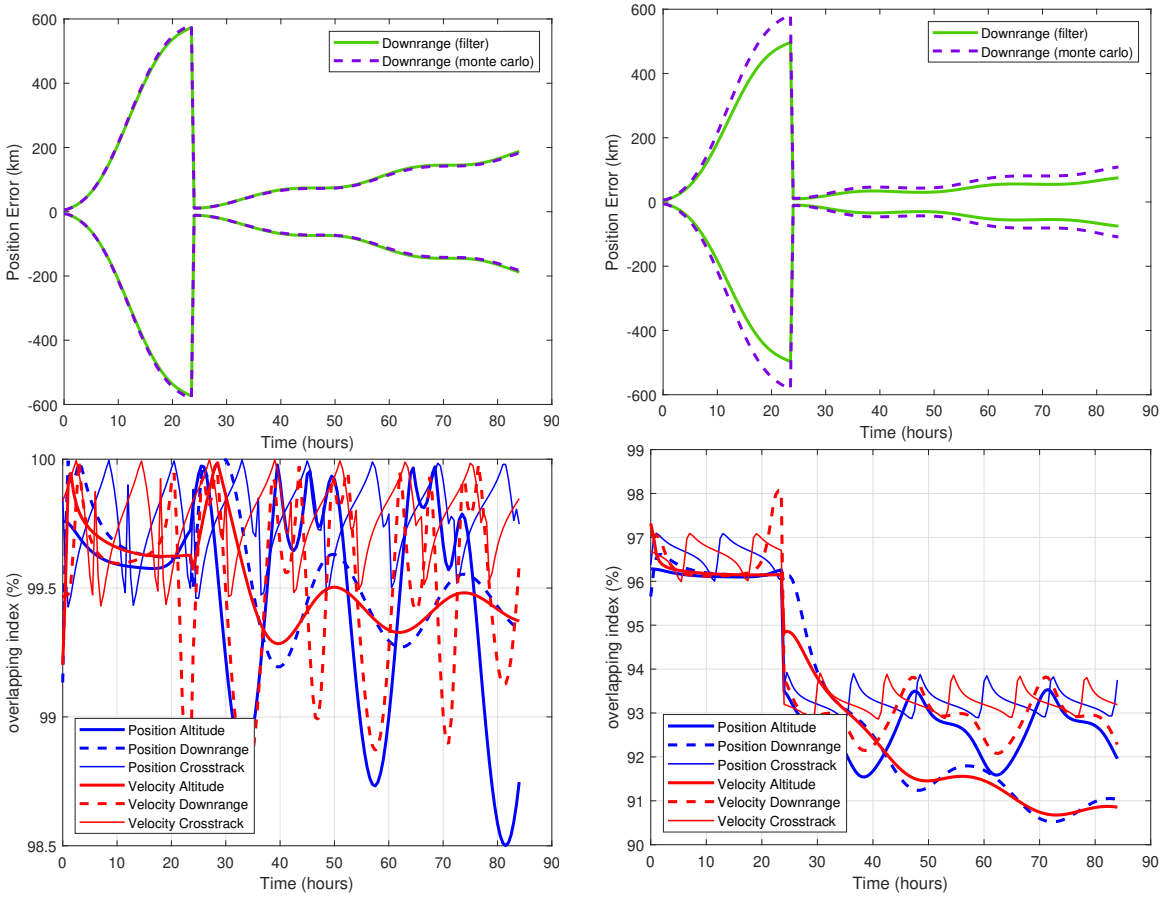
As illustrated in Figure 4.13, the ESBKF also tracks accurately when the OBS is at a higher altitude and the measurement is processed at 24 hours. The standard deviation values match very closely and covariance immediately following the measurement is small. Both the HKF and ESBKF are accurate, but the ESBKF is exceptionally more precise as well. Please refer to Figure F.50 for the results of 3σ values for all LVLH components.



(a) EKF (b) UKF

Fig. 4.14: 3σ and η EKF and UKF Measurement at 24 Hours, OBS 3 Only

The EKF and UKF are accurately tracking the RSO using a highly inclined OBS configuration. It was expected that the OBS 3 configuration, with an inclination angle of 30°, would yield a more precise post-measurement update. This is due to the higher inclination having more out-of-plane observability when compared to the other OBS configurations that are nearly coplanar with the RSO. The results shown in Figure 4.14 demonstrate that no significant improvement in filter performance was achieved by having more out-of-plane observability. Please refer to Figures F.3 and F.27 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.15: 3σ and η HKF and GMM Measurement at 24 Hours, OBS 3 Only

Figure 4.15 depicts the HKF and GMM under the scenario of a measurement obtained at 24 hours from an OBS with a 30° inclination. Both filter algorithms are tracking the RSO accurately. Please refer to Figures F.15 and F.39 for the results of 3σ values for all LVLH components.

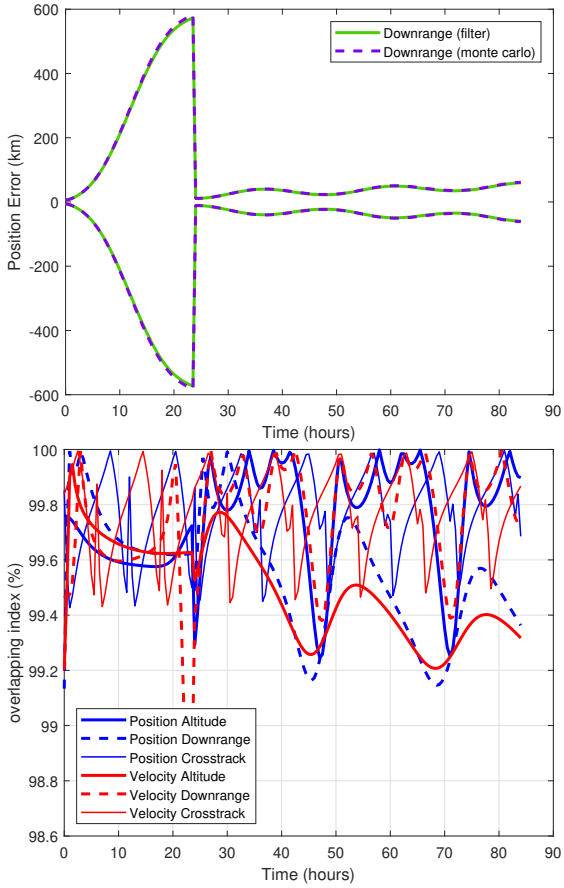
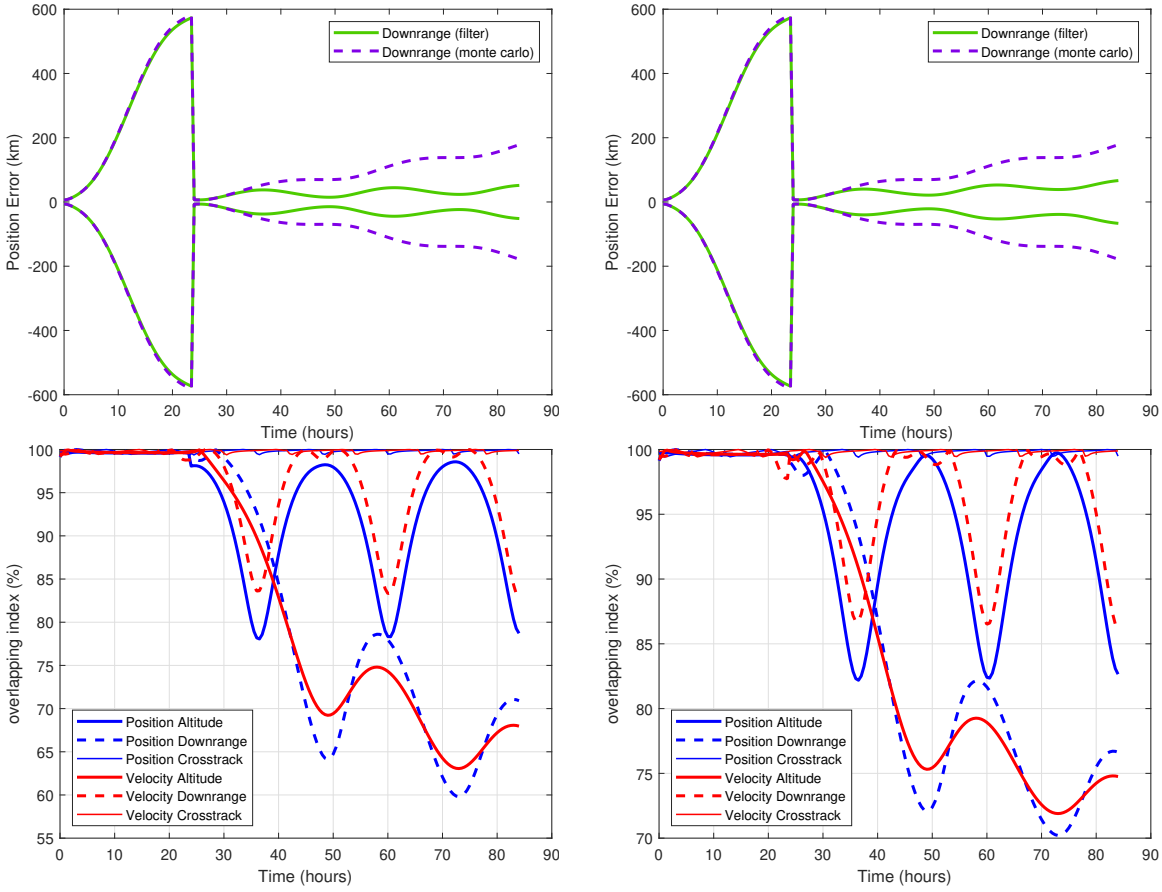


Fig. 4.16: 3σ and η ESBKF Measurement at 24 Hours, OBS 3 Only

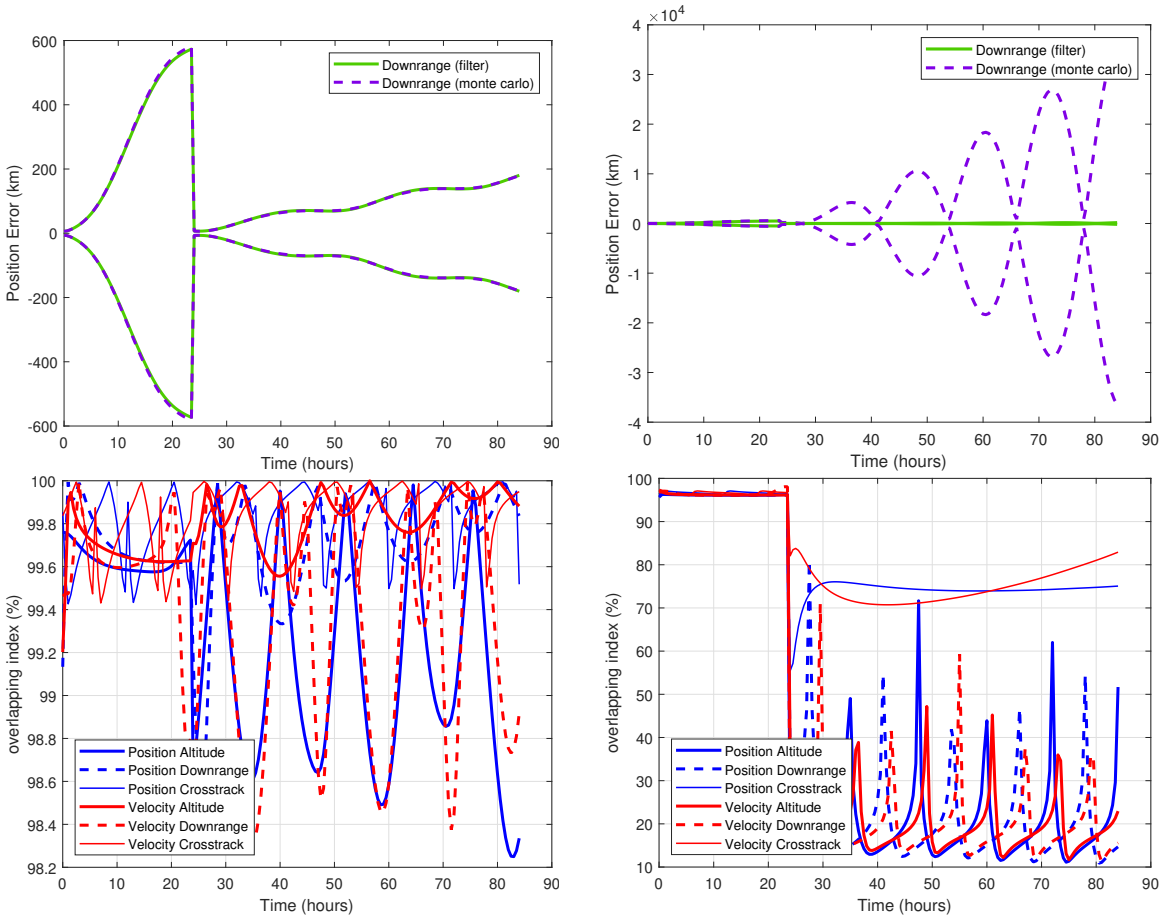
The ESBKF tracks the RSO accurately when the OBS is highly inclined, as illustrated in Figure 4.16. There is no significant change in ESBKF performance based any of the OBS configurations. Please refer to Figure F.51 for the results of 3σ values for all LVLH components.



(a) EKF (b) UKF

Fig. 4.17: 3σ and η EKF and UKF Measurement at 24 Hours, All Observations

As shown in Figure 4.17, the EKF falls below the 64% threshold for η as therefore divergent. It was expected that combining the OBS 1, OBS 2, and OBS 3 configurations would have resulted in better accuracy, even for the EKF, however, the results proved otherwise. Please refer to Figures F.4 and F.28 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.18: 3σ and η HKF and GMM Measurement at 24 Hours, All Observations

Figure 4.18 illustrates that the HKF is tracking the RSO accurately but the GMM falls well-below the threshold on 64% for the η. When using the GMM, if only one OBS is updating the RSO, the Kalman Gain moves the estimate to a location that, when used to evaluate the new weighting factor, gives a numerically stable result. However, when all three OBSs are updating the RSO, the Kalman Gain shifts the estimate to a location where all weighting factors cannot be handled numerically. This contradicts the intuition that having more measurements should improve observability and would thus be even more accurate. The conclusion is that the GMM is convergent only when one OBS is used. Please refer to Figures F.16 and F.40 for the results of 3σ values for all LVLH components.

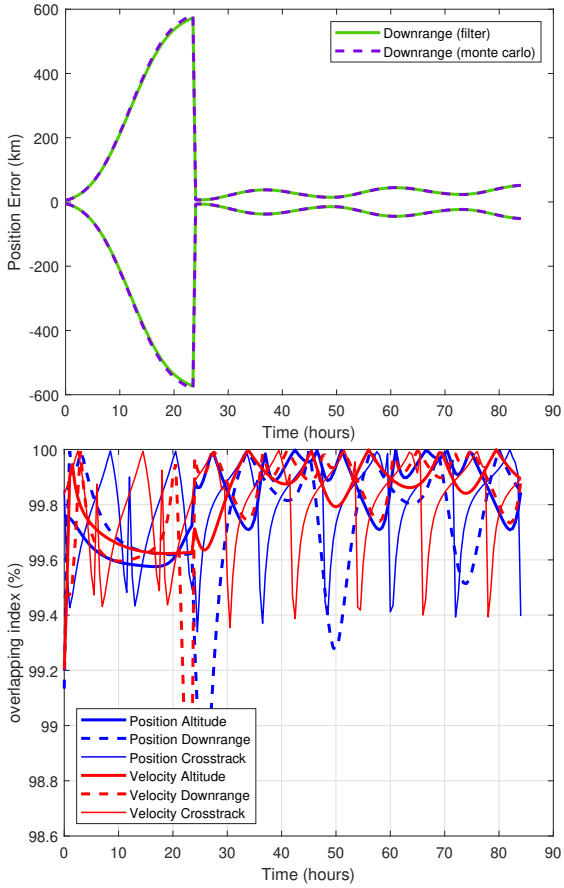
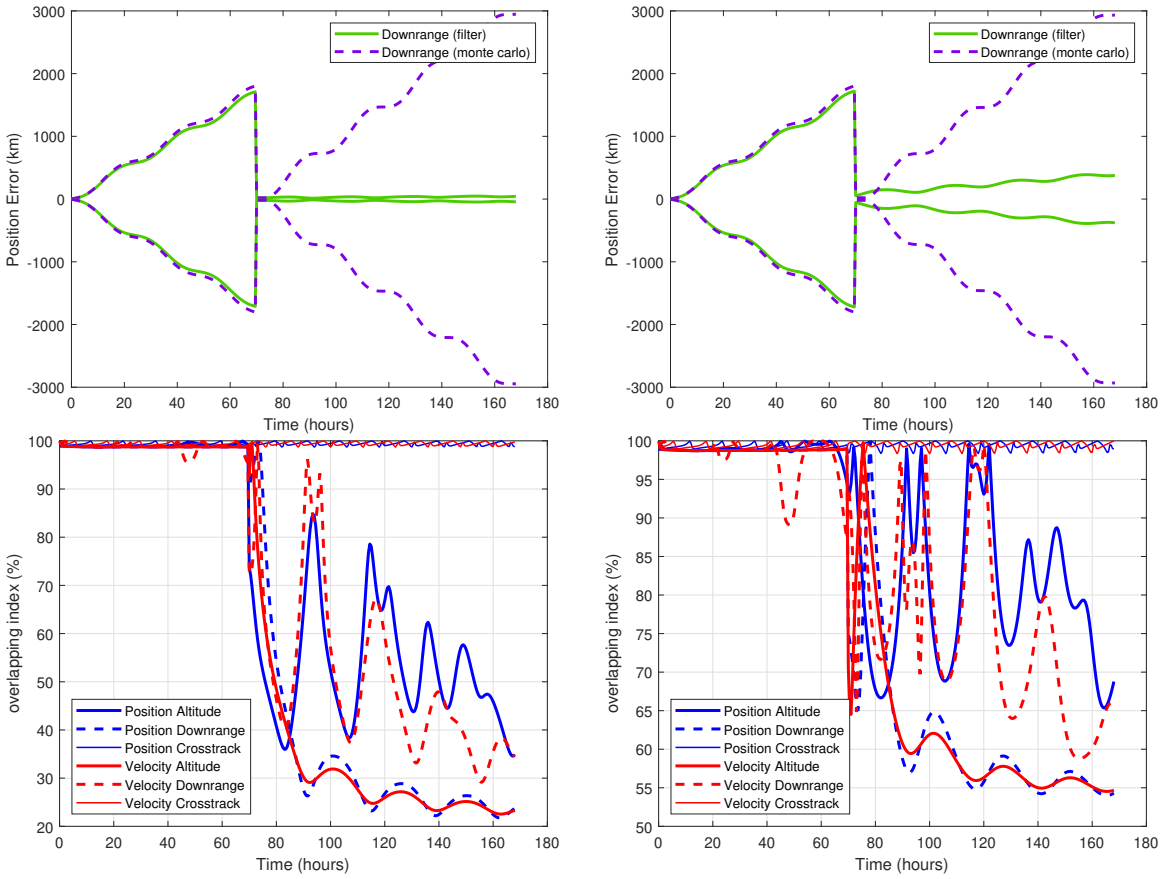


Fig. 4.19: 3σ and η ESBKF Measurement at 24 Hours, All Observations

The ESBKF tracks the RSO accurately when all OBS configurations are processed simultaneously, as illustrated in Figure 4.19. There is no significant change in ESBKF performance based any of the OBS configurations, either individually or combined. Please refer to Figure F.52 for the results of 3σ values for all LVLH components.

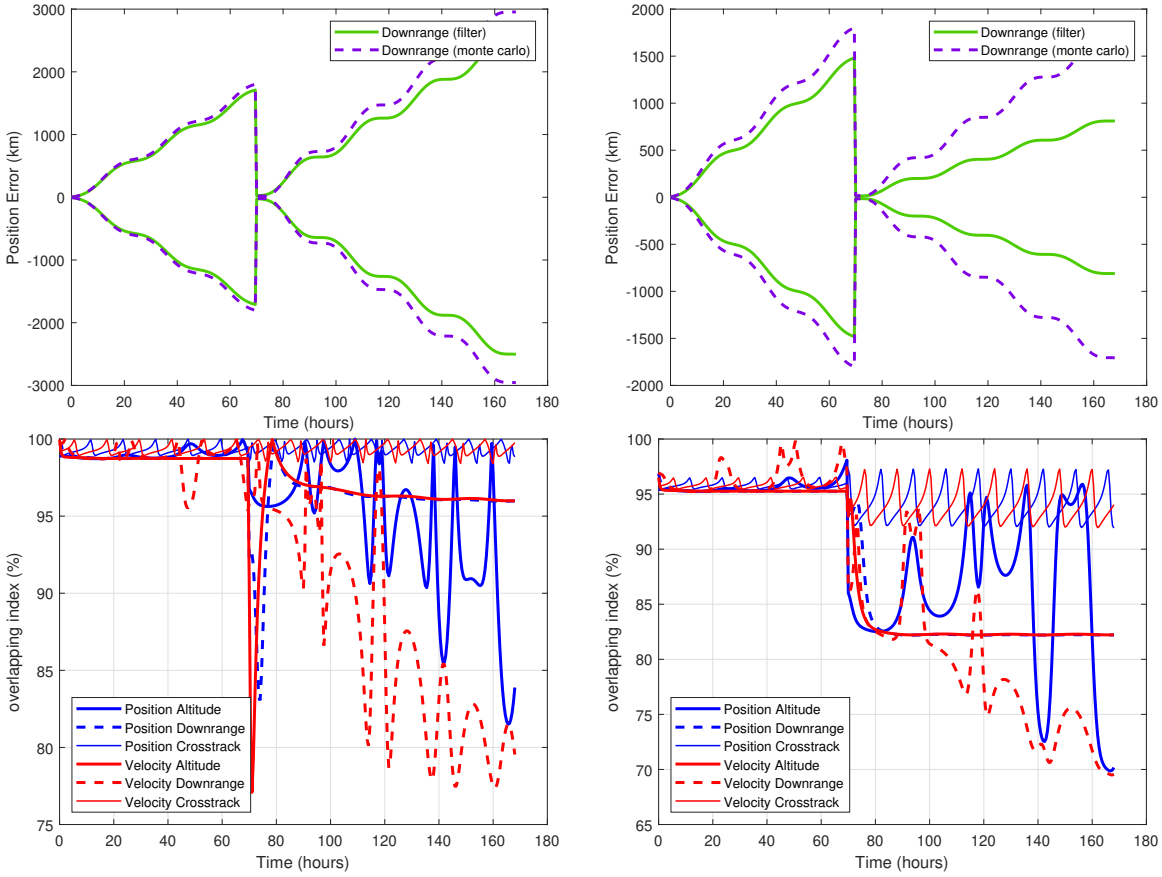
4.4.2 Measurement at 70 Hours



(a) EKF (b) UKF

Fig. 4.20: 3σ and η EKF and UKF Measurement at 70 Hours, OBS 1 Only

The EKF falls well-below the threshold for η and becomes divergent when the time of measurement is 70 hours and using the OBS 1 configuration. The HKF also falls below 64% for η, but not as severely as the EKF. Both filters are divergent as depicted in Figure 4.20. Please refer to Figures F.5 and F.29 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.21: 3σ and η HKF and GMM Measurement at 70 Hours, OBS 1 Only

The HKF and GMM both maintain accuracy at a measurement time of 70 hours and using the OBS 1 configuration. The accuracy of both filters has been reduced significantly, when compared to scenario of the measurement time of 24 hours, as shown in Figure 4.21. Please refer to Figures F.17 and F.41 for the results of 3σ values for all LVLH components.

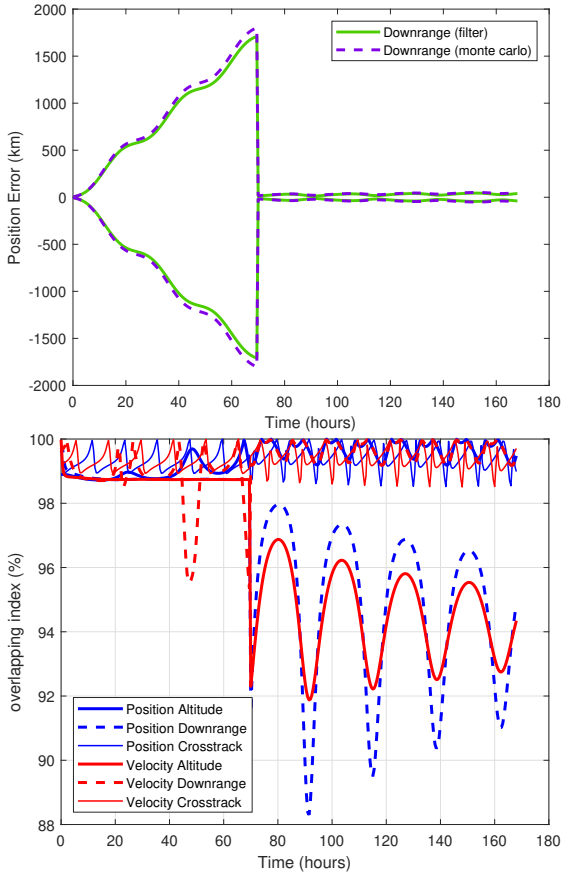
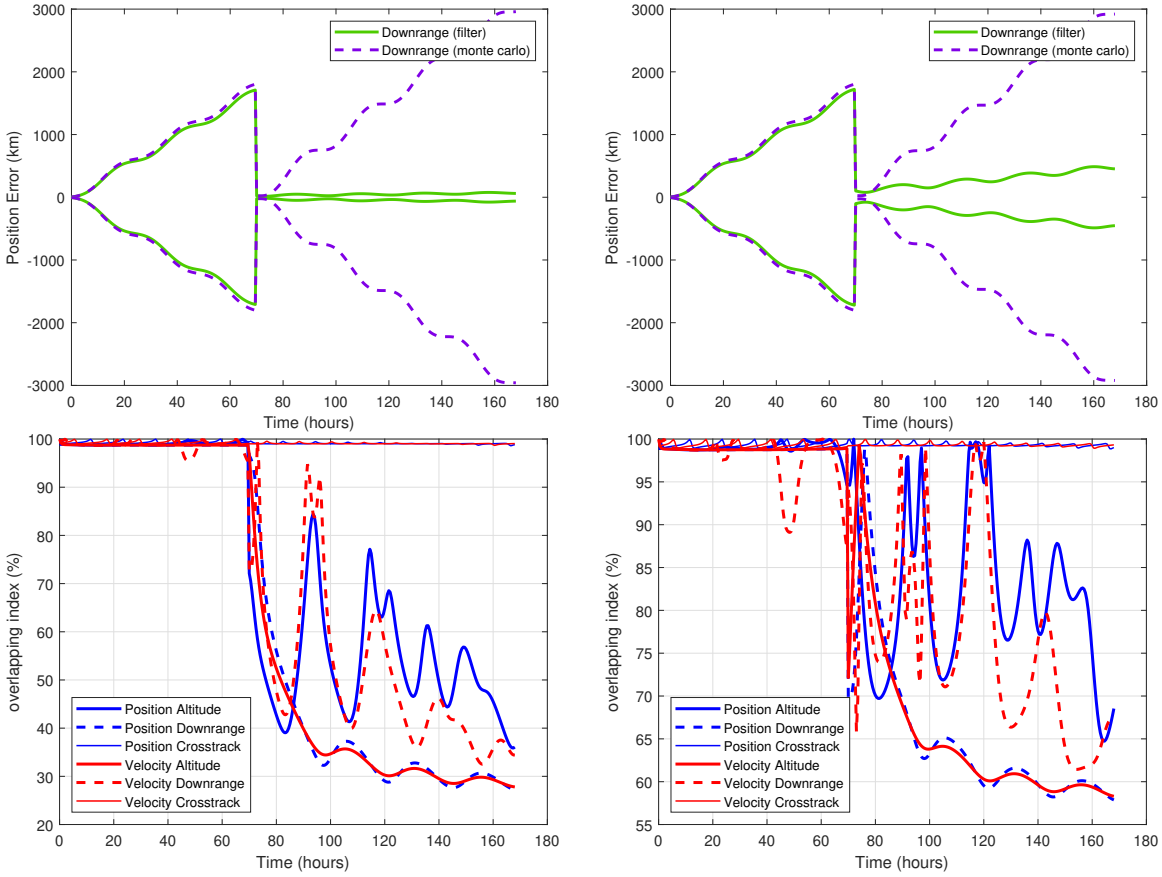


Fig. 4.22: 3σ and η ESBKF Measurement at 70 Hours, OBS 1 Only

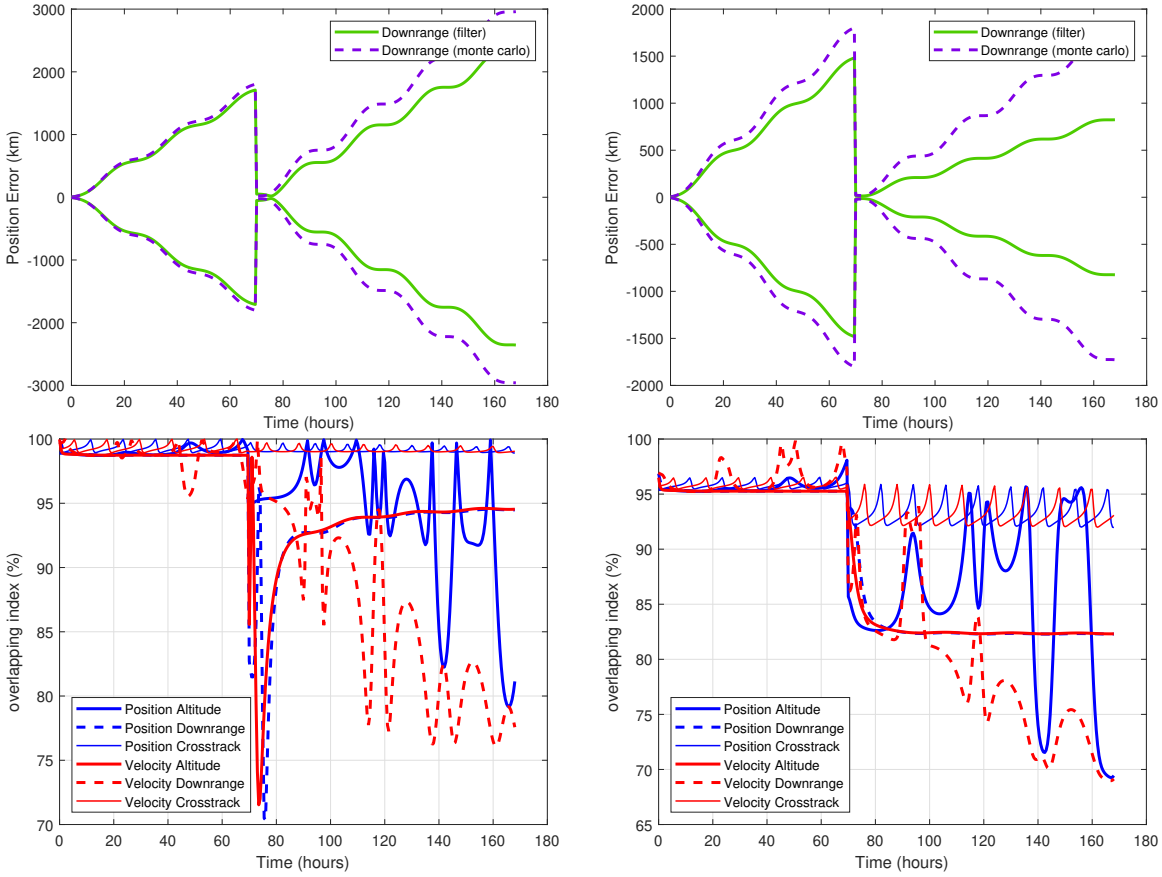
Figure 4.22 illustrates that the ESBKF maintains exceptional accuracy at the measurement time of 70 hours. The value of η dropped more compared to when the measurement was processed at 24 hours, but then it rebounds after the measurement update. Please refer to Figure F.53 for the results of 3σ values for all LVLH components.



(a) EKF (b) UKF

Fig. 4.23: 3σ and η EKF and UKF Measurement at 70 Hours, OBS 2 Only

The performance of both the EKF and the UKF is similar when the OBS 2 configuration is used as the OBS 1 configuration. Both filters diverge, but the UKF is not as severe as the EKF, as shown in Figure 4.23. Please refer to Figures F.6 and F.30 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.24: 3σ and η HKF and GMM Measurement at 70 Hours, OBS 2 Only

Both the HKF and GMM accurately track the RSO when the OBS 2 configuration is used at a measurement time of 70 hours, as depicted in Figure 4.24. When compared with the EKF and UKF, the HKF and GMM maintain better accuracy. Please refer to Figures F.18 and F.42 for the results of 3σ values for all LVLH components.

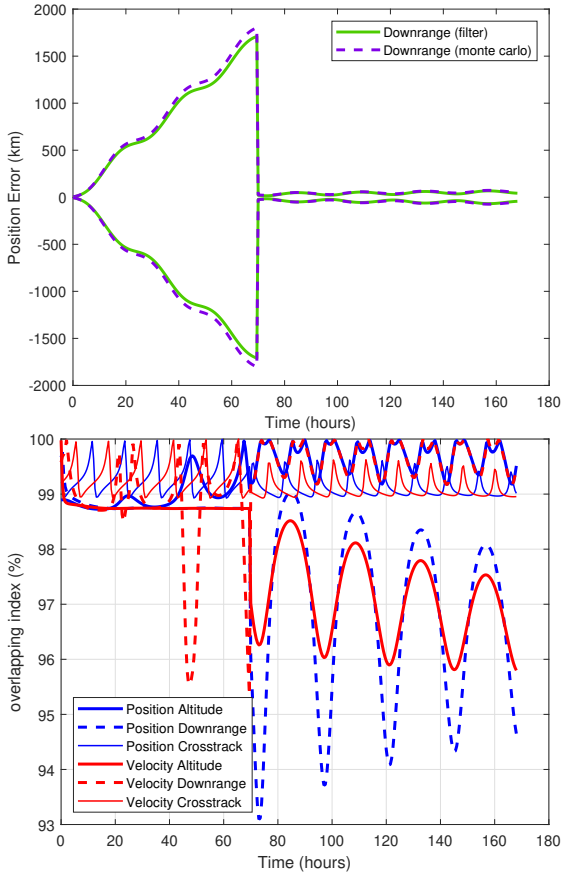
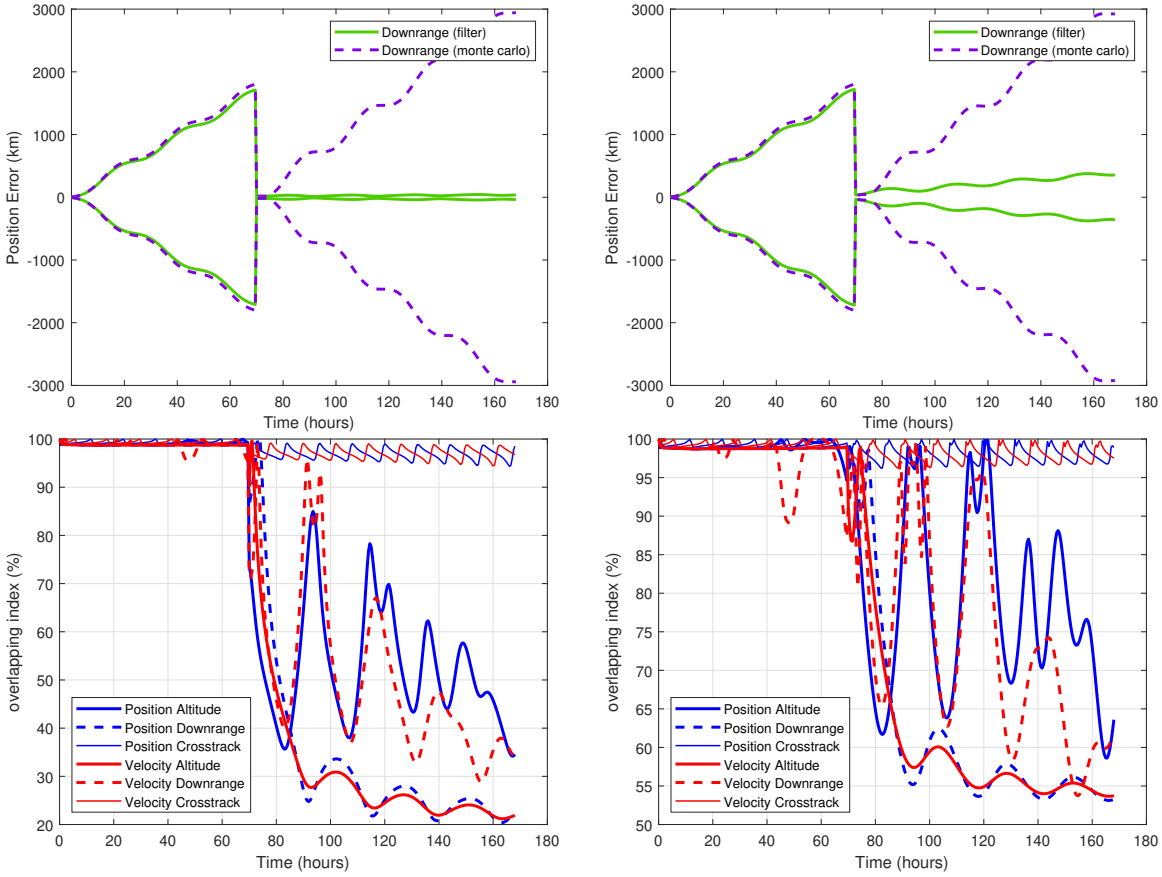


Fig. 4.25: 3σ and η ESBKF Measurement at 70 Hours, OBS 2 Only

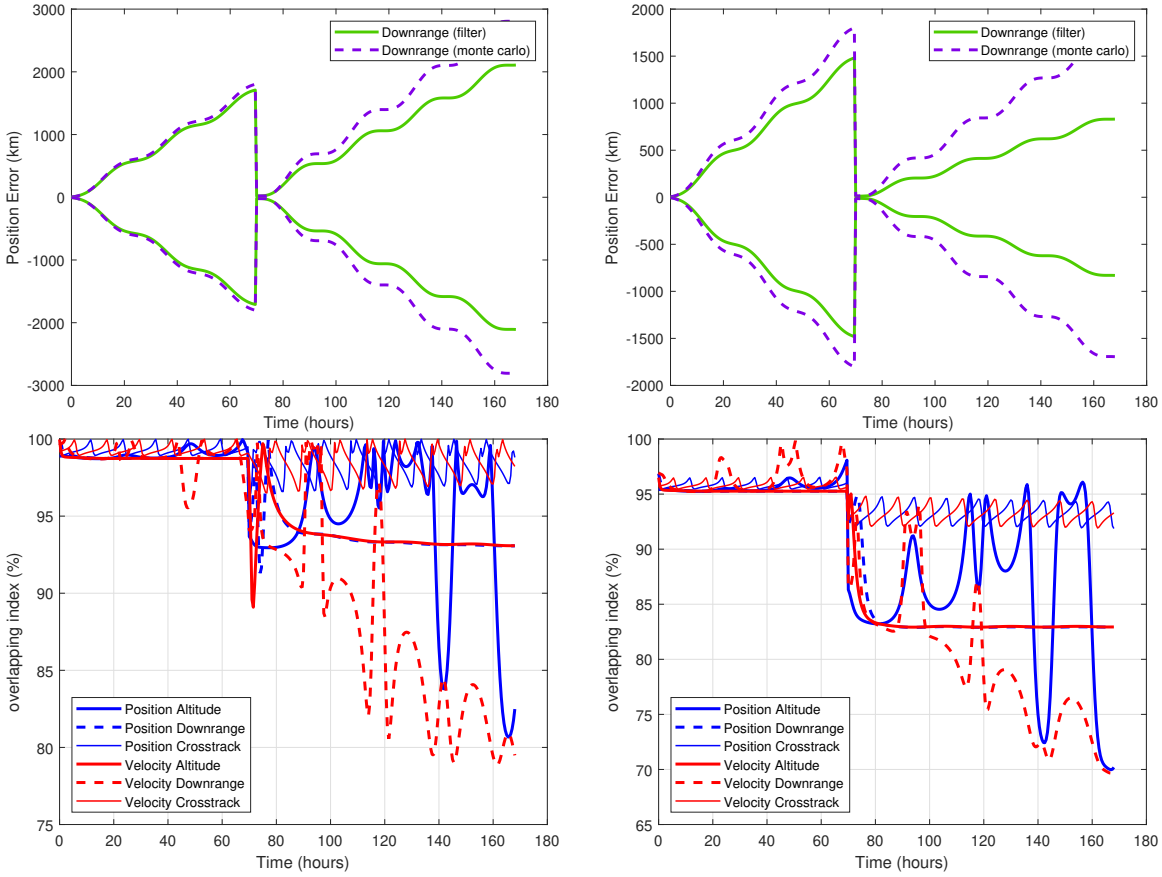
As illustrated in Figure 4.25, the ESBKF has exceptional accuracy when compared with the other filter algorithms in the scenario of using the OBS 2 configuration at a measurement time of 70 hours. Please refer to Figure F.54 for the results of 3σ values for all LVLH components.



(a) EKF (b) UKF

Fig. 4.26: 3σ and η EKF and UKF Measurement at 70 Hours, OBS 3 Only

The performance of both the EKF and the UKF is similar when the OBS 1 or OBS 2 configurations are used. Both filters diverge, but the UKF is not as severe as the EKF, as shown in Figure 4.26. Please refer to Figures F.7 and F.31 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.27: 3σ and η HKF and GMM Measurement at 70 Hours, OBS 3 Only

Both the HKF and GMM accurately track the RSO when the OBS 3 configuration is used at a measurement time of 70 hours, as depicted in Figure 4.27. As seen in previous results, the HKF and GMM maintain accuracy longer than the EKF and UKF. Please refer to Figures F.19 and F.43 for the results of 3σ values for all LVLH components.

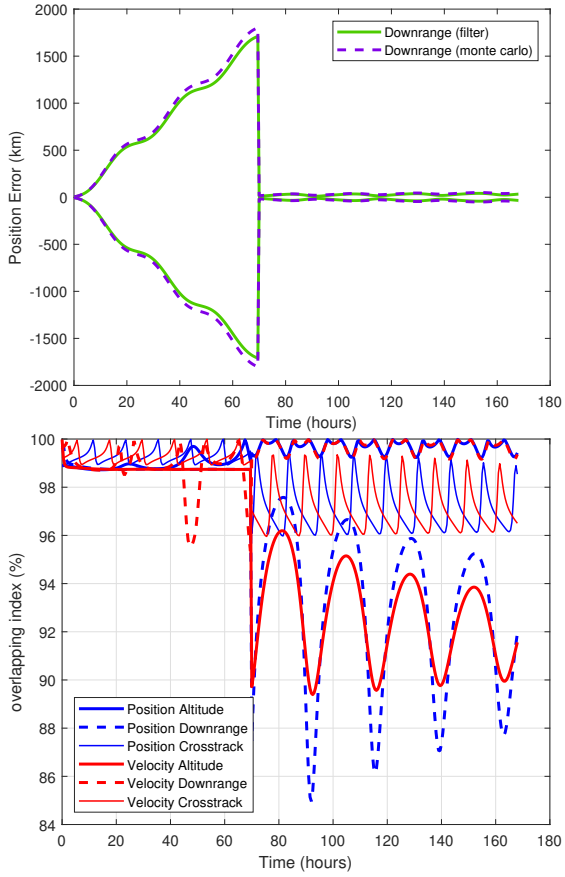
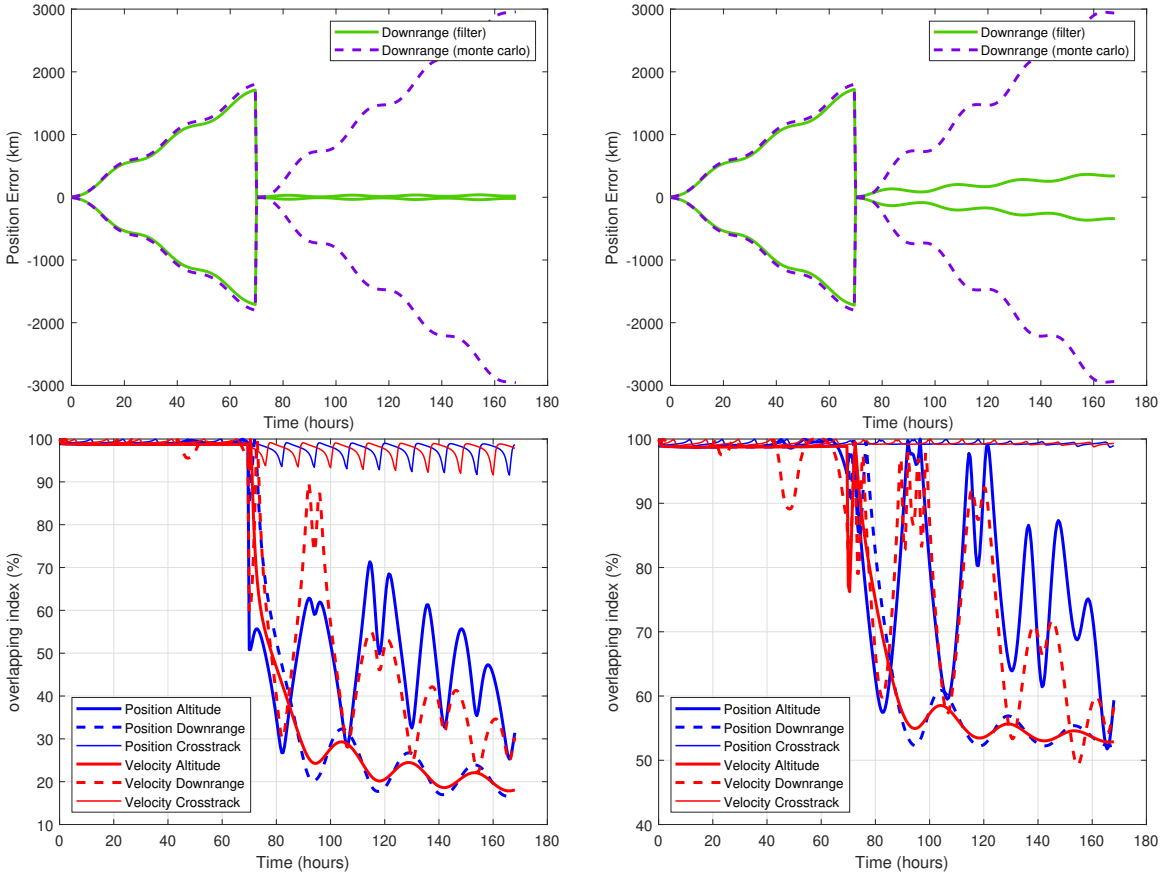


Fig. 4.28: 3σ and η ESBKF Measurement at 70 Hours, OBS 3 Only

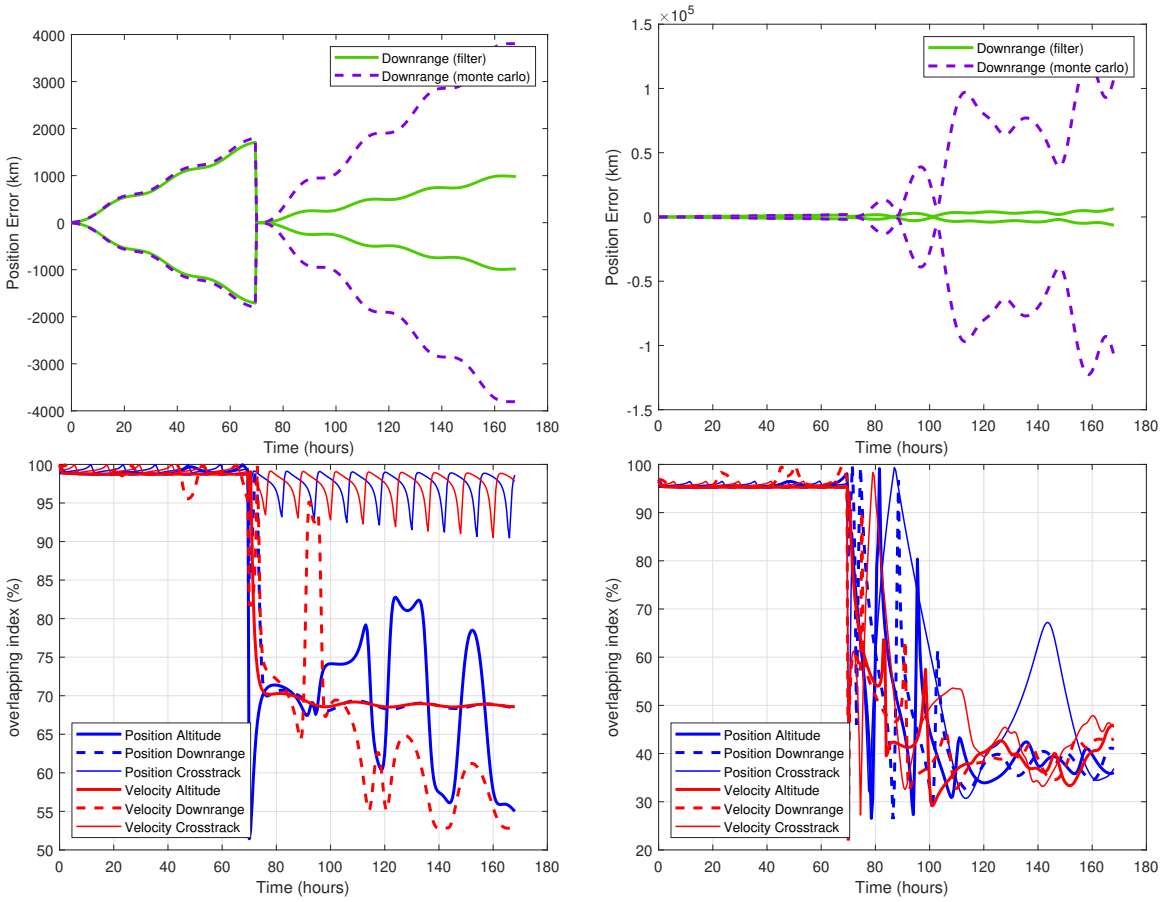
Figure 4.28 illustrates that the ESBKF maintains exceptional accuracy when using the OBS 3 configuration at the measurement time of 70 hours. The value of η dropped more compared to when the measurement was processed at 24 hours. Please refer to Figure F.55 for the results of 3σ values for all LVLH components.



(a) EKF (b) UKF

Fig. 4.29: 3σ and η EKF and UKF Measurement at 70 Hours, All Observations

Figure 4.29 illustrates that the EKF and UKF both diverge when all three of the OBS configurations are combined and processed simultaneously. Please refer to Figures F.8 and F.32 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.30: 3σ and η HKF and GMM Measurement at 70 Hours, All Observations

The HKF and GMM both diverge when using all OBS configurations simultaneously. The behavior of the GMM when all OBS configurations are used remains the same as in the other scenarios. Numerical issues arise when updating the weighting factors in this scenario. This is depicted in Figure 4.30. Please refer to Figures F.20 and F.44 for the results of 3σ values for all LVLH components.

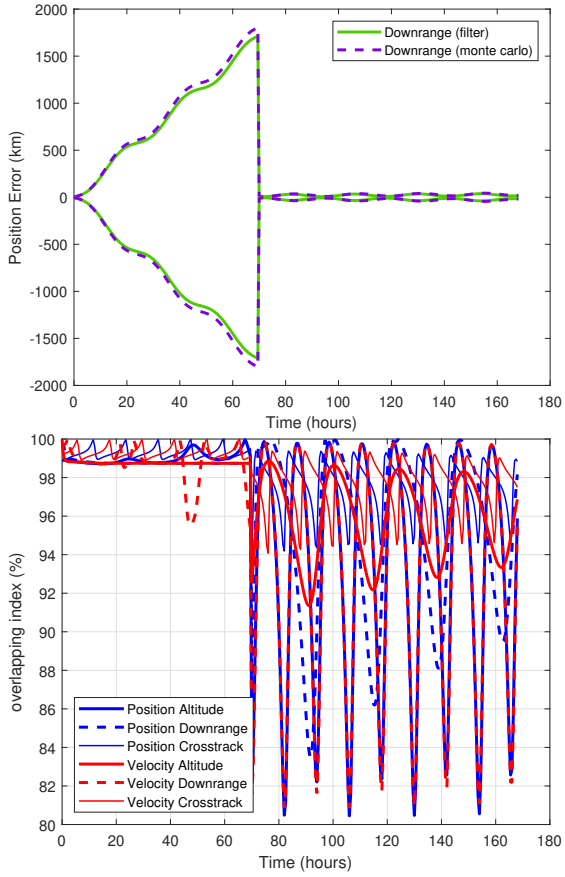
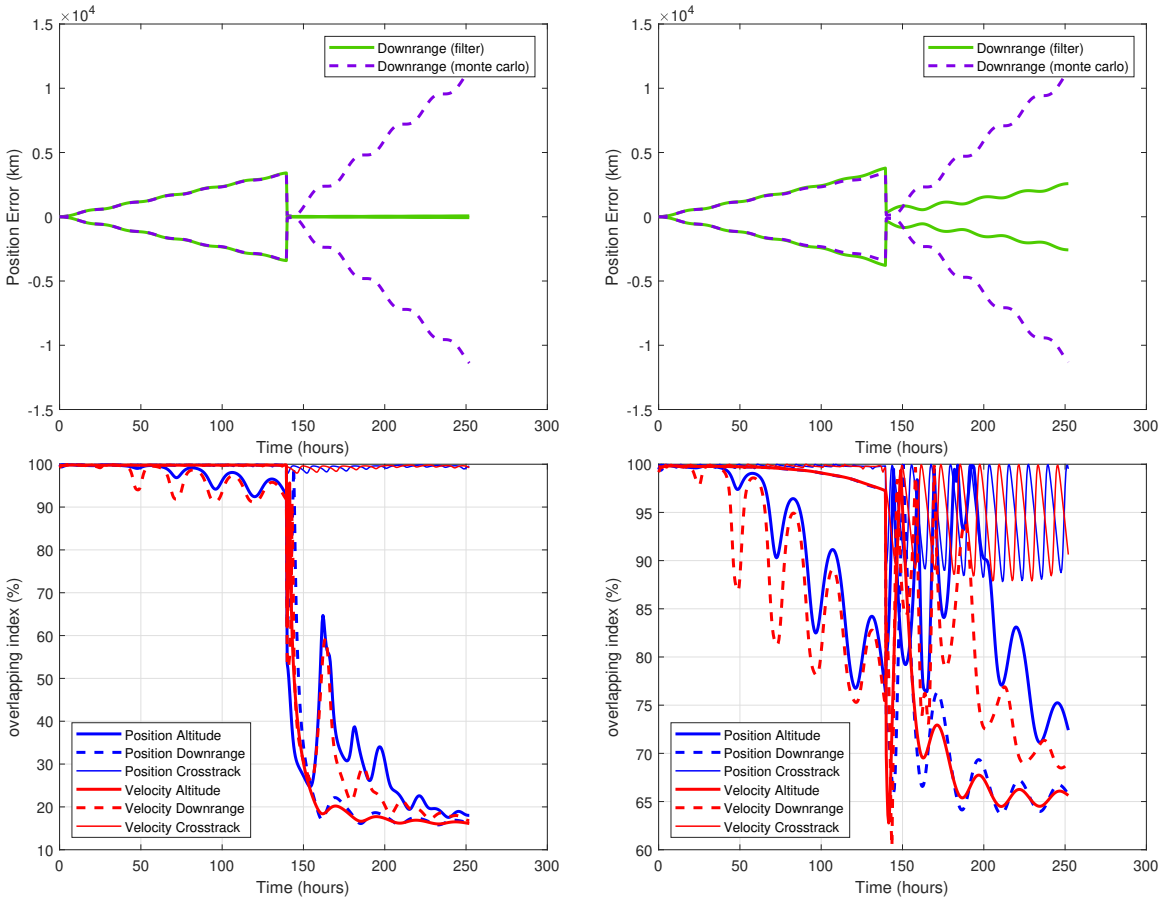


Fig. 4.31: 3σ and η ESBKF Measurement at 70 Hours, All Observations

The ESBKF was the only filter algorithm that maintained accuracy where all OBS configurations were processed simultaneously at the measurement time of 70 hours, as illustrated in Figure 4.31. The dependability of the ESBKF for maintaining accuracy when measurements are taken at 70 hours is demonstrated. Please refer to Figure F.56 for the results of 3σ values for all LVLH components.

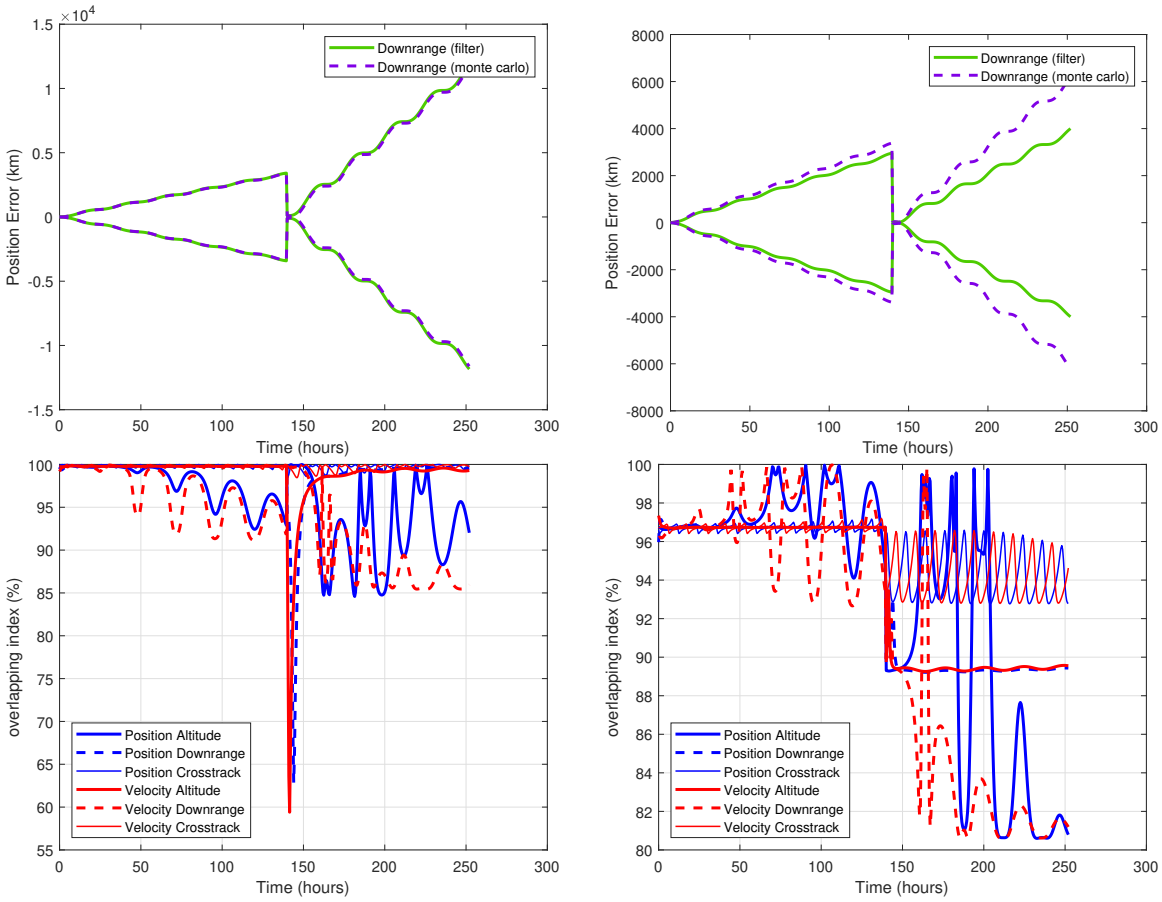
4.4.3 Measurement at 140 Hours



(a) EKF (b) UKF

Fig. 4.32: 3σ and η EKF and UKF Measurement at 140 Hours, OBS 1 Only

When measurements were processed at 140 hours using the OBS 1 configuration, both the EKF and UKF diverged, as depicted in Figure 4.32. The accuracy of the UKF, even though it eventually diverged, was much higher than that of the EKF. Please refer to Figures F.9 and F.33 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.33: 3σ and η HKF and GMM Measurement at 140 Hours, OBS 1 Only

As shown in Figure 4.33, the HKF drops below the threshold for η immediately following the measurement update, however, it rebounded to a high level of accuracy. There is not much confidence in this filter in this scenario of using the OBS 1 configuration at the measurement time of 140 hours. The GMM maintained accuracy. Please refer to Figures F.21 and F.45 for the results of 3σ values for all LVLH components.

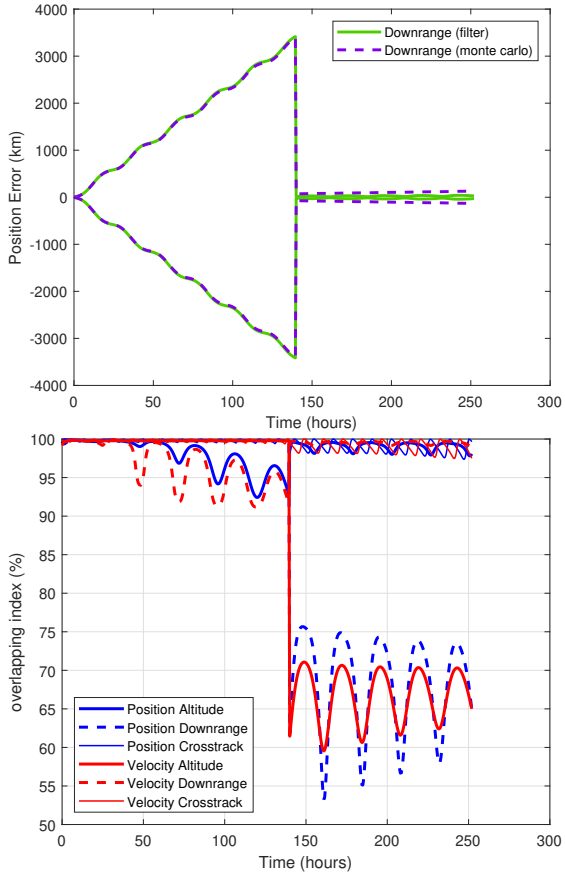
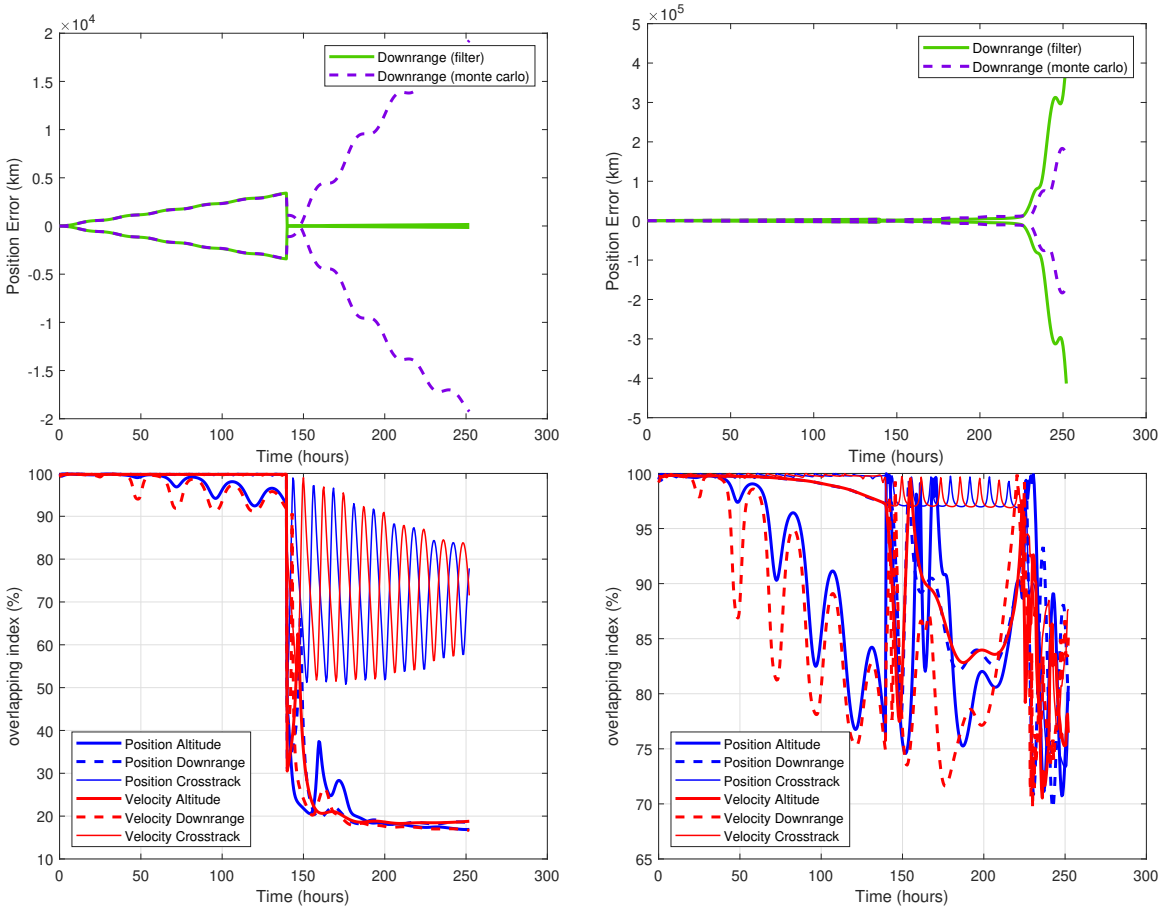


Fig. 4.34: 3σ and η ESBKF Measurement at 140 Hours, OBS 1 Only

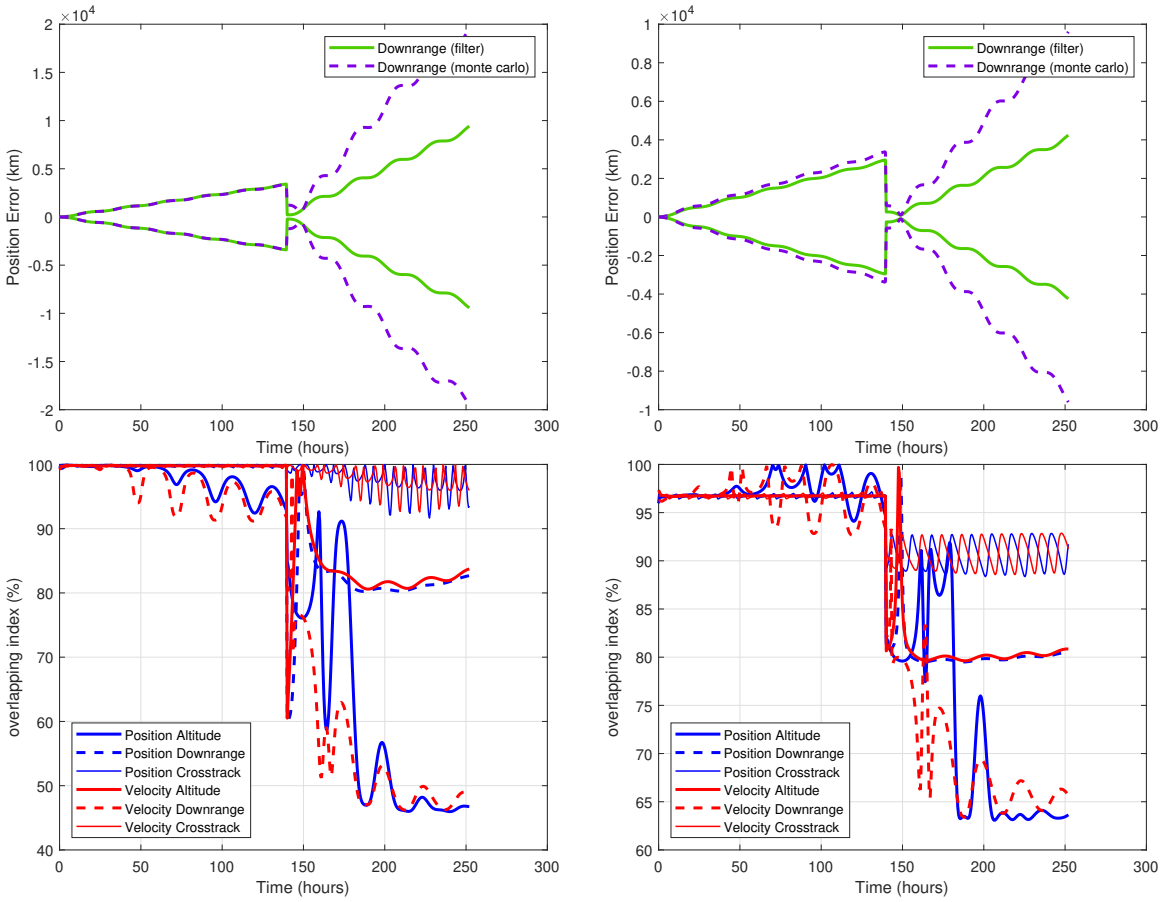
As shown in Figure 4.34, the ESBKF becomes divergent at the measurement time of 140 hours and using the OBS 1 configuration. This can be attributed to the fact that while the ESBKF removes the issue of not having a PDF that is Gaussian for the state vector when performing a Kalman update, there remains nonlinearity in the measurement model. This conclusion is corroborated from the PDFs in Figure 4.7 in which the location of the RSO after the measurement update is offset from the truth and has a smaller covariance. An item of future work will be to develop a technique into the Step-Back Kalman filter methodology that would account for nonlinear measurements. This may be investigating an unscented form of the SBKF, for example. Please refer to Figure F.57 for the results of 3σ values for all LVLH components.



(a) EKF (b) UKF

Fig. 4.35: 3σ and η EKF and UKF Measurement at 140 Hours, OBS 2 Only

Figure 4.35 illustrates that the EKF diverges while the UKF maintains accuracy based on η values. The scenario is using the OBS 2 configuration at a measurement time of 140 hours. The EKF is clearly divergent when looking at both the 3σ and η plots, but the UKF appears to have a large surge in the 3σ plot which could lead to future divergence and yet the η plot for the UKF demonstrates that it maintained accuracy. Please refer to Figures F.10 and F.34 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.36: 3σ and η HKF and GMM Measurement at 140 Hours, OBS 2 Only

Figure 4.36 shows that both the HKF and GMM diverge, however, the GMM barely falls below the 64% threshold. The scenario uses the OBS 2 configuration at the measurement time of 140 hours. Please refer to Figures F.22 and F.46 for the results of 3σ values for all LVLH components.

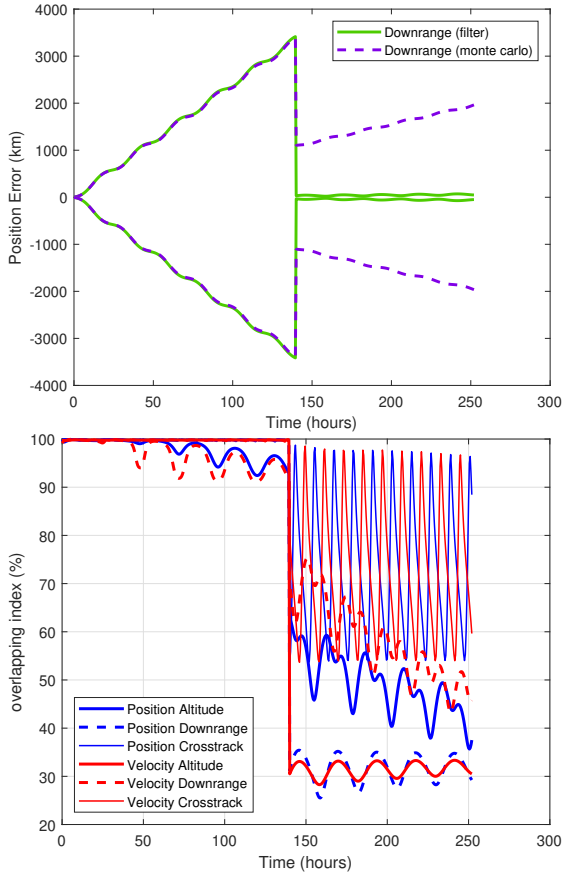
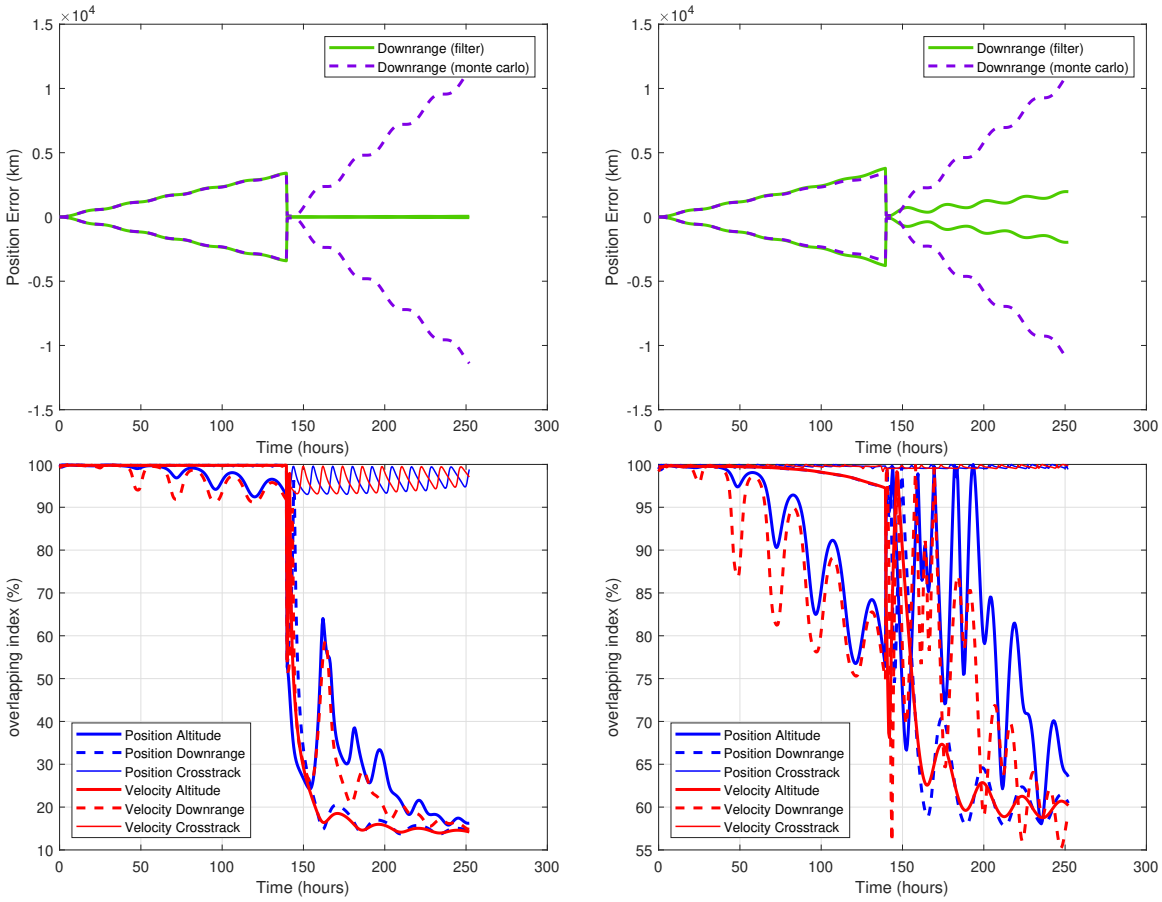


Fig. 4.37: 3σ and η ESBKF Measurement at 140 Hours, OBS 2 Only

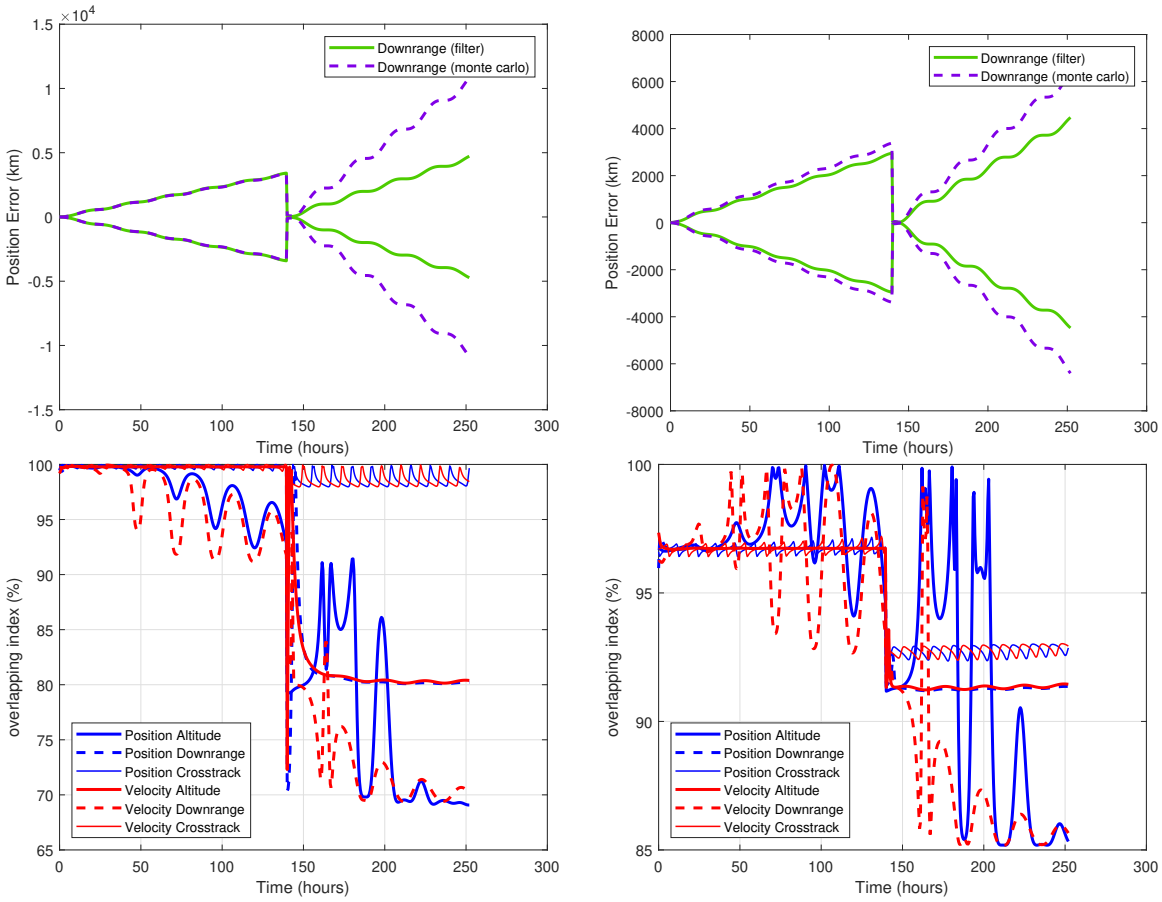
As shown in Figure 4.37, the ESBKF becomes divergent at the measurement time of 140 hours and using the OBS 2 configuration. The ESBKF does not maintain accuracy at the measurement time of 140 hours. Please refer to Figure F.58 for the results of 3σ values for all LVLH components.



(a) EKF (b) UKF

Fig. 4.38: 3σ and η EKF and UKF Measurement at 140 Hours, OBS 3 Only

Figure 4.38 illustrates that both the EKF and UKF diverge in the scenario using OBS 3 at the measurement time of 140 hours. The reason why the UKF does not have as low a value for η as the EKF is that, at the time of measurement, the UKF estimates a larger covariance matrix, which can compensate for the error of the mean value estimate. Please refer to Figures F.11 and F.35 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.39: 3σ and η HKF and GMM Measurement at 140 Hours, OBS 3 Only

Figure 4.39 illustrates that the HKF and GMM both track the RSO accurately. The scenario uses the OBS 3 configuration at the measurement time of 140 hours. Please refer to Figures F.23 and F.47 for the results of 3σ values for all LVLH components.

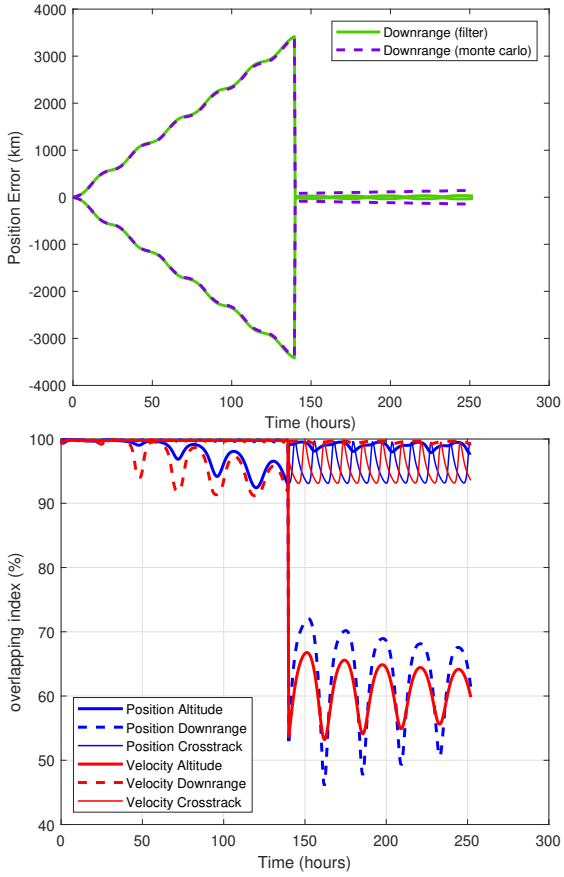
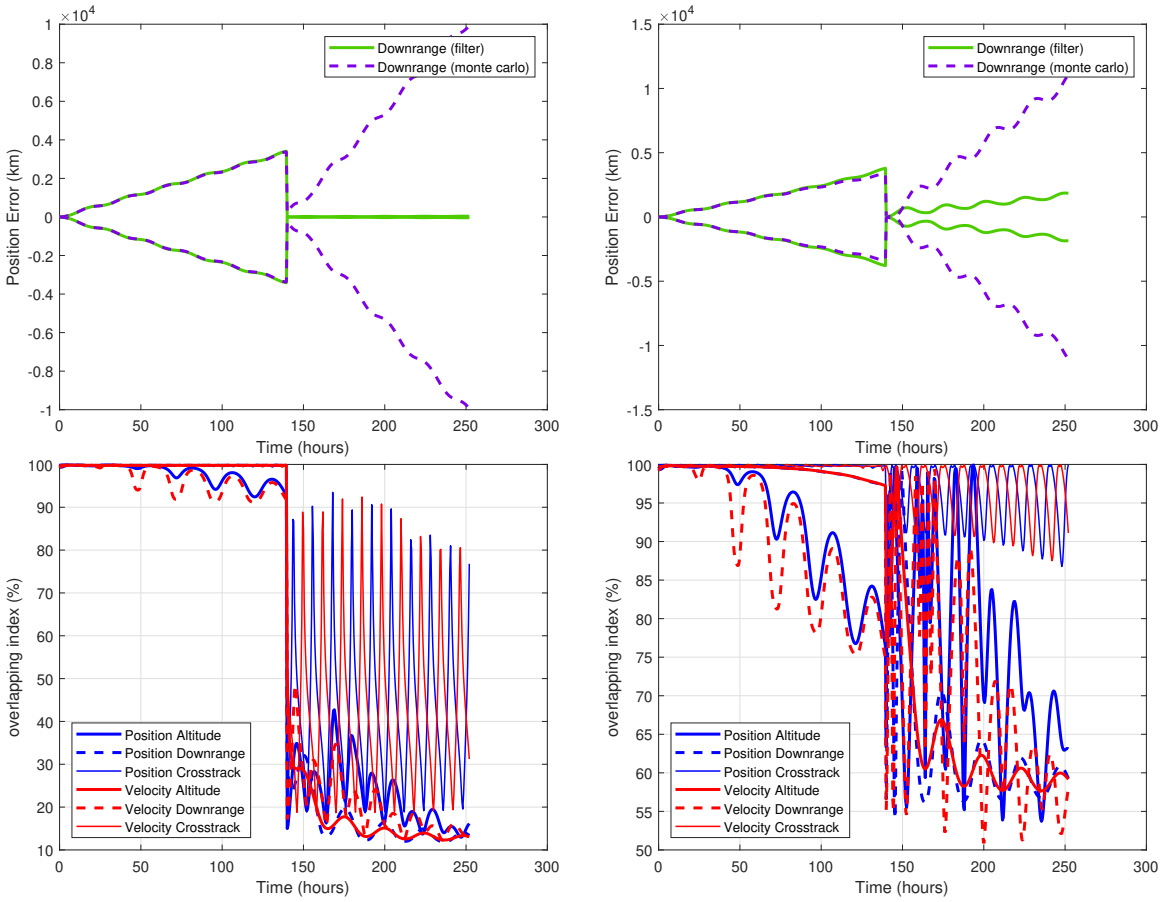


Fig. 4.40: 3σ and η ESBKF Measurement at 140 Hours, OBS 3 Only

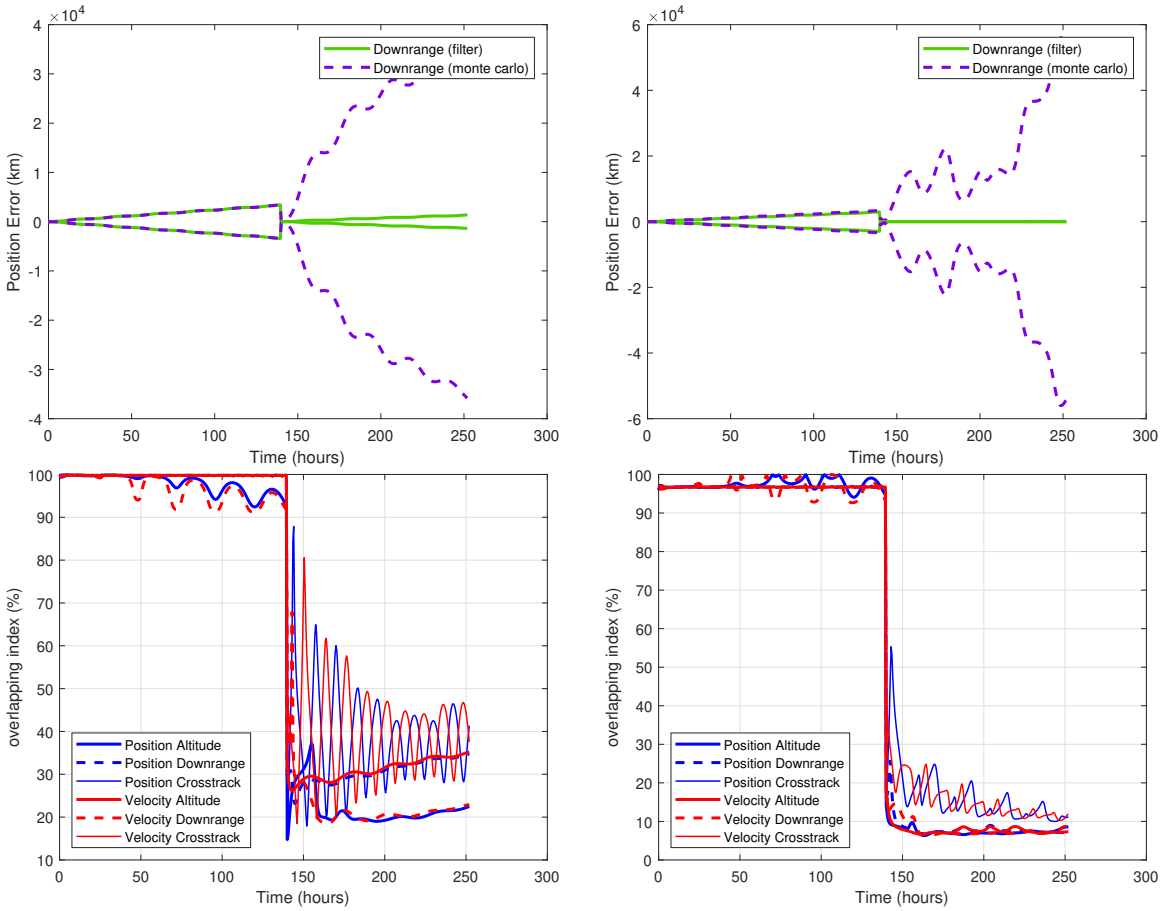
The ESBKF diverges, as shown in Figure 4.40. This is a case where other filter algorithms outperformed the ESBKF. Changing the OBS configurations makes a difference for some algorithms in certain situations. Please refer to Figure F.59 for the results of 3σ values for all LVLH components.



(a) EKF (b) UKF

Fig. 4.41: 3σ and η EKF and UKF Measurement at 140 Hours, All Observations

Figure 4.41 illustrates that both the EKF and UKF diverge in the scenario using all three OBS configurations at the measurement time of 140 hours. Please refer to Figures F.12 and F.36 for the results of 3σ values for all LVLH components.



(a) HKF (b) GMM

Fig. 4.42: 3σ and η HKF and GMM Measurement at 140 Hours, All Observations

Figure 4.42 depicts that both the HKF and GMM diverge severely. The scenario uses all OBS configurations simultaneously at the measurement time of 140 hours. Please refer to Figures F.24 and F.48 for the results of 3σ values for all LVLH components.

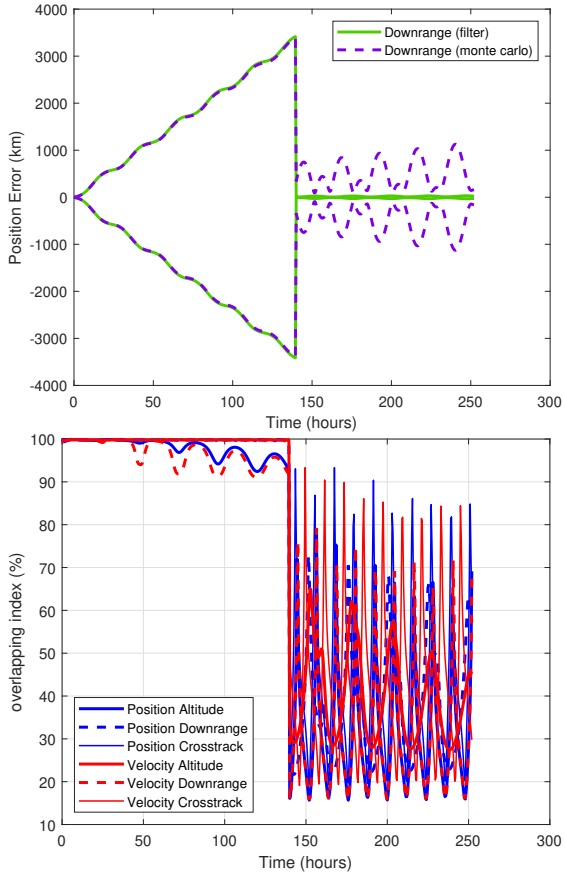


Fig. 4.43: 3σ and η ESBKF Measurement at 140 Hours, All Observations

The ESBKF diverges, as shown in Figure 4.43. This a case where all filter algorithms diverged. The scenario uses all OBS configurations simultaneously at the measurement time of 140 hours. Please refer to Figure F.60 for the results of 3σ values for all LVLH components.

4.5 Computation Time Comparisons

Table 4.1 and Table 4.2 contain the computational speed of the filtering algorithms quantified as a ratio with the computational speed of the EKF. Table 4.1 and Table 4.2 do not match well for the HKF. The HKF uses sample statistics for the mean and covariance used in the Kalman update by propagating 1,000 samples directly from the time of last measurement to the current time using Keplerian dynamics. The particles always used

Keplerian propagation. Numerical integration is a much more computationally burdensome task than a Keplerian propagation, therefore when full dynamics were used, the HKF is practically on par with the EKF but when the filter uses a Kepler routine to propagate, then the time it takes to propagate is on par with the particle propagation to acquire sample statistics and thus the increase ratio in Table 4.2.

Table 4.1: Computation Time Comparison with Full Dynamics

Filter Algorithm	Ratio of Computation Time with EKF
EKF	1
HKF	1.08
UKF	6.66
GMM (2 nodes per dimension)	54.64
ESBKF	1.46

Table 4.2: Computation Time Comparison with Keplerian Dynamics

Filter Algorithm	Ratio of Computation Time with EKF
EKF	1
HKF	3.29
UKF	6.01
GMM (2 nodes per dimension)	44.24
GMM (3 nodes per dimension)	476.01
GMM (4 nodes per dimension)	2669.93
ESBKF	1.02

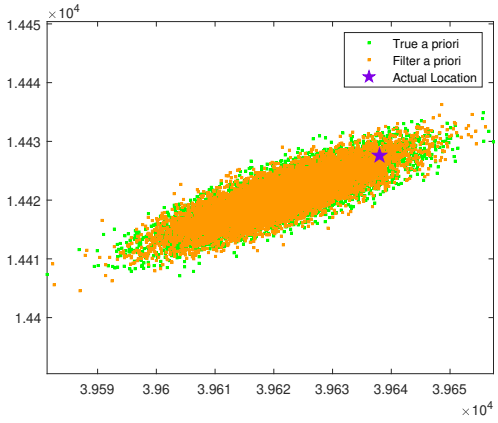
The computation time required to integrate the position and velocity mean value with the covariance matrix augmented together as a combined state vector in the EKF is a factor of 1.72 times greater than integrating the position and velocity alone. It was anticipated

that the factor would be approximately 7 since the covariance matrix has 36 elements and the mean value has 6 elements. The result of 1.72 indicates the extra function calls with the differential equations of motion for the Jacobian matrices adds almost double the computation time required when performing a Runge-Kutta integration scheme. In the UKF, each particle is individually integrated in a for-loop which calls the Runge-Kutta integration function 13 separate times per time step rather than calling only once with an augmented state vector. It was found that the computational time for integrating all 13 particles is a factor of 7.84 times greater than integrating the augmented state vector in the EKF. This ratio of 7.84 differs slightly from the values of 6.66 and 6.01 in Tables 4.1 and E.1, respectively, since both EKF and UKF also have to process measurement updates in different formulations. An additional burden of computation in the UKF is obtaining a square-root decomposition of the covariance matrix, typically Cholesky decomposition, when acquiring new sigma points at each time step [17], [36]. This is why a square-root formulation of the UKF may also be used which has the same performance in terms of accuracy as the UKF but is formulated to increase computational speed [37]. An area of interest for future research is to apply the SRUKF to this simulation and compare the computation speed.

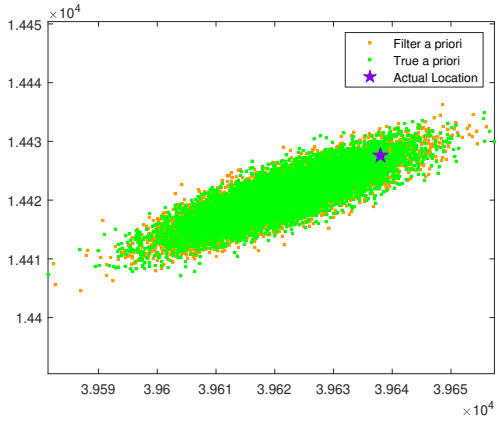
4.6 Closed Skew-Normal Unscented Kalman Filter

From the beginning of this research, it was found that using the assumption of a Gaussian distribution for the state under nonlinear behavior is what leads to post-measurement update divergence in a filter. After the UKF and HKF were simulated and performance assessed, it became apparent that simply trying to obtain a better mean value and covariance matrix when applying a Kalman measurement update is insufficient since it attempts to solve the true problem of skewness accumulating in the PDF by enlarging the covariance to cover the spread. A larger Gaussian is not the same as a thinner skewed PDF and thus these filters had only moderate improvements compared to the EKF. The heuristic then became to investigate the prospects of a filter which tracks the third moment, skewness. A filter algorithm of this form had already been developed, the CSNKF. It follows a Kalman filter

methodology but with other parameters than the mean and covariance. A skewness tensor is not tracked directly in the same way the EKF tracks the first two moments directly, but rather there are several parameters which shape the distribution to have skewness. It was a concern that tracking the skewness would have hindered the computational speed since a third order tensor would have $6^3 = 216$ elements. This CSNKF became the focus of the research to become a resolution to the filter divergence problem. The CSNKF, however, had a linearized formulation following with the EKF and if the PDF is initialized as a Gaussian, then linear propagation would retain the Gaussian geometry and skewness would not accrue. However, there is an unscented form of the CSNKF, the CSNUKF, which would allow one to initialize sigma points based on a Gaussian distribution at $t = 0$ but then nonlinear dynamics would passively induce skewness by propagation of the sigma points. At each time steps the sigma points would be gathered and the CSN parameters would be parsed and skewness would be accrued. Before the CSNUKF was applied to the full simulation, a proof-of-concept check was performed by comparing CSN parameters propagated to a known time when skewness is present following CSNUKF and sampling from the CSN distrbution to see how well it matched the PDF extracted from propagating particles following the orbital dynamics. As seen in Figures 4.44 and 4.45, this did not match well. The PDF remained close to Gaussian and so the skewness did not accrue passively through dynamical evolution. Skewness was then inserted directly into the PDF at $t = 0$ to determine whether seeding skewness immediately would assist in accruing skewness passively. Figures 4.46 and 4.47 demonstrate that this did not improve the problem. Another drawback of the CSNUKF is the computational requirements for performance evaluation. The computation time for performing the MCMC with Metropolis-Hastings using only 100 Monte Carlo samples and a markov-chain of 20 runs translates to a factor of about 650 times greater in computation time compared to the EKF just in parsing the mean value and covariance matrix from the CSN. As a result of these drawbacks, the CSNUKF was removed as the focus of this research to find a reasonable solution to the divergence problem of the EKF.

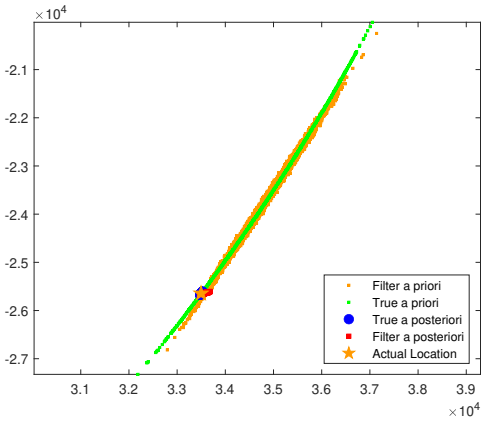


(a) Position XY plane filter above

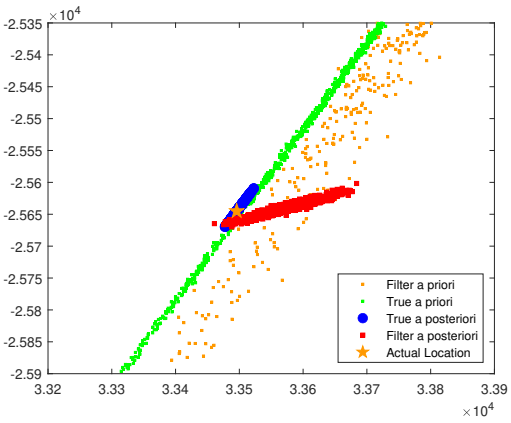


(b) Position XY plane filter below

Fig. 4.44: Closed Skew-Norm Unscented Kalman filter density function at 0 hours (**A**)

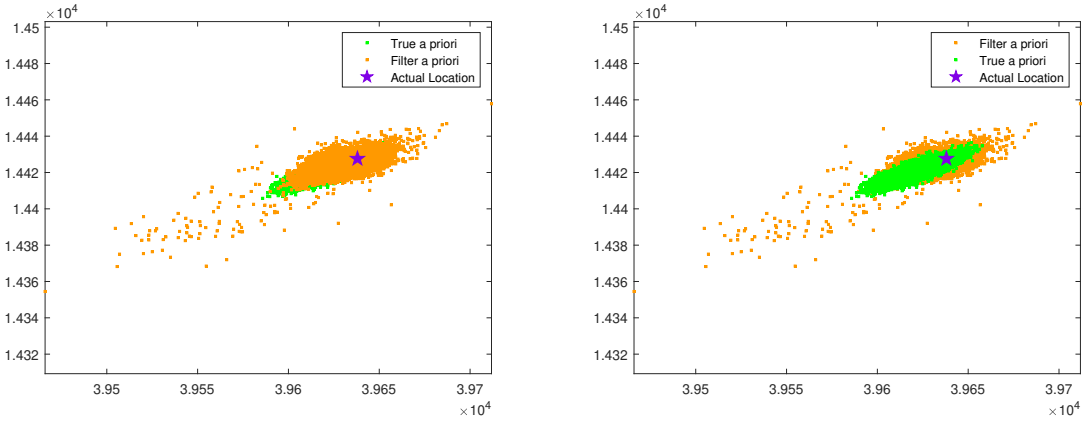


(a) Position XY plane zoomed out view



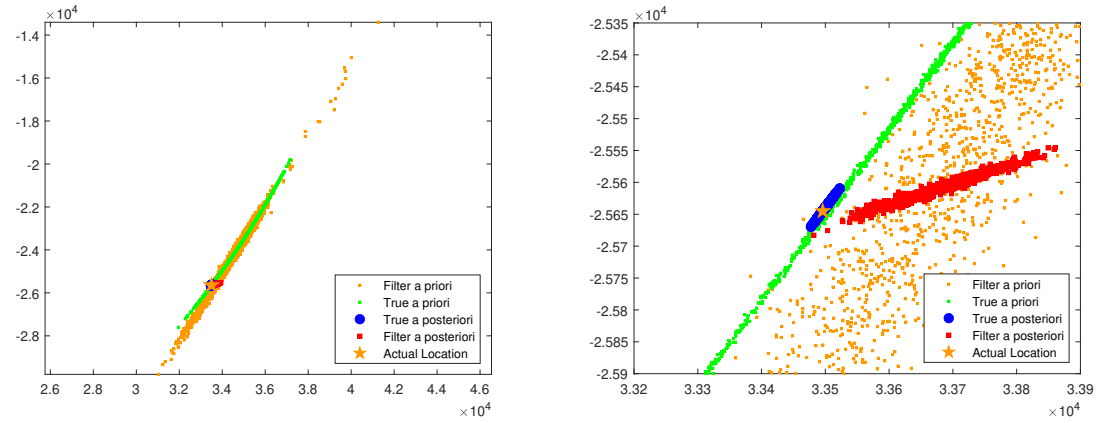
(b) Position XY plane zoomed in view

Fig. 4.45: Closed Skew-Norm Unscented Kalman filter density function at 140 hours (**A**)



(a) Position XY plane filter above (b) Position XY plane filter below

Fig. 4.46: Closed Skew-Norm Unscented Kalman filter density function at 0 hours (**B**)



(a) Position XY plane zoomed out view (b) Position XY plane zoomed in view

Fig. 4.47: Closed Skew-Norm Unscented Kalman filter density function at 140 hours (**B**)

4.7 Summary

The following list presents the performance of all of the filter algorithms,

- When the time when a measurement update is processed is 24 hours, the EKF diverges when all OBS configurations are combined, the UKF performs accurately for all scenarios, the HKF performs accurately for all scenarios, the GMM performs ac-

curately for all scenarios except for when all OBS configurations are combined, and the ESBKF performs accurately for all scenarios.

- When the measurement is being acquired at 70 hours, the EKF diverges in all scenarios, the UKF diverges in all scenarios, the HKF performs accurately for all scenarios, the GMM performs accurately for all scenarios except for when all OBS configurations are combined, and the ESBKF performs accurately for all scenarios.

- When the measurement is being acquired at 140 hours, the EKF diverges in all scenarios, the UKF performs accurately only when OBS 2 is used, the HKF performs accurately only when OBS 3 is used, the GMM performs accurately when OBS 1 and OBS 3 are used but it diverges when OBS 2 and all OBS configurations are used, and the ESBKF diverges in all scenarios.

The CSNUKF filter, was not fully simulated to run 3σ and η plot analysis due to the enormous computation time required to numerically calculate the mean and covariance. This computational burden is too vast such that the CSNUKF would not be a feasible candidate to use in actual scenarios. However, scatter plot analysis was performed to give a comparison to how the CSNUKF compared with the other algorithms. It was the anticipated that the CSNUKF would have formed a skewed distribution that would adhere close to the geometry of the true distribution. Also, it was anticipated that using the Unscented form of the Closed Skew-Normal Kalman filter would have resulted in the filter probability distribution evolving to match somewhat closely to the true probability distribution in a passive manner. The sigma points would be forced to propagate through the nonlinear dynamics thereby placing them in locations such that the parameters defining the Closed Skew-Normal distribution would accurately represent a PDF that matches closely with the true PDF. It turns out that this was not the case.

CHAPTER 5

CONCLUSIONS

Tracking and surveillance of RSOs in orbit around Earth, particularly in GEO where communication satellites are most present, is critical to avoid orbital debris. Traditionally, the Extended Kalman filter has been used for nonlinear dynamic behavior filtering algorithms, but it has limitations based on the underlying assumptions of linearized dynamics and measurements, and Gaussian distributions for the state and measurement.

For short durations between measurements, the EKF handles the filtering process adequately in most situations, but once nonlinear effects become significant, the algorithm will place an updated state estimate in the incorrect location. This is due to the fact that a Gaussian approximation will not adequately approximate certain regions of the true distribution when skewness is present. The intersection of the measurement distribution and the actual state distribution is the region where the updated state distribution actually resides. The EKF will place the intersection in a region away from where it should reside, generating post-update divergence as the filter proceeds further in time.

A novel algorithm was developed in this research to avoid the problem of nonlinear dynamics entirely. This algorithm is referred to as the Step-Back filter, and its nonlinear extension is the Extended Step-Back Kalman filter. The foundational premise behind this algorithm is that when a measurement is available to process for an update of the state vector statistics, the update is applied to the last point in time when the state PDF is known to be well-approximated as a Gaussian. The points in time when this is known are immediately following a measurement update or at the start of the algorithm based on initial conditions. Therefore, when a measurement update is applied to that point in time rather than the current time, skewness present in the true distribution is no longer relevant, leaving nonlinear issues from the measurement model as the primary error source. The updated statistics, when propagated back to the current time, will be much closer to

the actual statistics.

Other filtering algorithms perform well to varying degrees, but some require a large quantity of computational time to achieve accuracy. It should be noted that the Step-Back Kalman filter is not limited to only the extended form, but future work could look into an unscented form, a Gaussian Mixture form, and other formulations that the Kalman filter currently applies.

REFERENCES

[1] Horwood, J. T., Poore, A. B., and Alfriend, K. T., "Orbit determination and data fusion in GEO," *Proc. of the 2011 AMOS Conference,(Wailea, HI)*, 2011.

[2] Andrews, B. and Geller, D., "Analysis of Angles-Only Hybrid Space-Based/Ground-Based Approach for Geosynchronous Orbit Catalog Maintenance," *2018 Space Flight Mechanics Meeting*, 2018, p. 0726.

[3] Ristic, B., Arulampalam, S., and Gordon, N., *Beyond the Kalman filter: Particle filters for tracking applications*, Artech house, 2003.

[4] Wan, E. A. and Van Der Merwe, R., "The unscented Kalman filter for nonlinear estimation," *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, Ieee, 2000, pp. 153–158.

[5] Horwood, J. T. and Poore, A. B., "Adaptive Gaussian sum filters for space surveillance," *IEEE transactions on automatic control*, Vol. 56, No. 8, 2011, pp. 1777–1790.

[6] Bruck, R. F. and Copley, R. H., "GEODSS present configuration and potential," Tech. rep., AIR FORCE SPACE COMMAND MAUI HI 21ST OPERATIONS GROUP, 2014.

[7] Vallado, D. A. and Griesbach, J. D., "Simulating space surveillance networks," *Paper AAS 11-580 presented at the AAS/AIAA Astrodynamics Specialist Conference. July*, 2011.

[8] Sharma, J., Stokes, G. H., von Braun, C., Zollinger, G., and Wiseman, A. J., "Toward operational space-based space surveillance," *Lincoln Laboratory Journal*, Vol. 13, No. 2, 2002, pp. 309–334.

[9] Fujimoto, K. and Scheeres, D. J., "Non-linear Bayesian orbit determination based on the generalized admissible region," *2012 15th International Conference on Information Fusion*, IEEE, 2012, pp. 2043–2049.

[10] Maybeck, P. S., *Stochastic models, estimation, and control*, Academic press, 1982.

[11] Woodburn, J. and Tanygin, S., "Detection of non-linearity effects during orbit estimation," *Paper AAS*, 2010, pp. 10–239.

[12] Julier, S. J. and Uhlmann, J. K., "New extension of the Kalman filter to nonlinear systems," *Signal processing, sensor fusion, and target recognition VI*, Vol. 3068, Spie, 1997, pp. 182–193.

[13] Mashiku, A., Garrison, J., and Carpenter, J. R., "Statistical Orbit Determination using the Particle Filter for incorporating Non-Gaussian Uncertainties." *AIAA/AAS Astrodynamics Specialist Conference*, 2012, p. 5063.

[14] Pardal, P. C. P. M., Kuga, H. K., and De Moraes, R. V., "The particle filter sample impoverishment problem in the orbit determination application," *Mathematical Problems in Engineering*, Vol. 2015, 2015.

[15] Van Der Merwe, R., Wan, E., and Julier, S., "Sigma-point Kalman filters for nonlinear estimation and sensor-fusion: Applications to integrated navigation," *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2004, p. 5120.

[16] Menegaz, H. M., Ishihara, J. Y., Borges, G. A., and Vargas, A. N., "A systematization of the unscented Kalman filter theory," *IEEE Transactions on automatic control*, Vol. 60, No. 10, 2015, pp. 2583–2598.

[17] Van Der Merwe, R. and Wan, E. A., "The square-root unscented Kalman filter for state and parameter-estimation," *2001 IEEE international conference on acoustics, speech, and signal processing. Proceedings (Cat. No. 01CH37221)*, Vol. 6, IEEE, 2001, pp. 3461–3464.

[18] Psiaki, M. L., "Backward-smoothing extended Kalman filter," *Journal of guidance, control, and dynamics*, Vol. 28, No. 5, 2005, pp. 885–894.

[19] Taplin, R. H., "The reverse kalman filter," *Communications in Statistics-Theory and Methods*, Vol. 27, No. 10, 1998, pp. 2547–2558.

[20] Roonizi, A. K., "Kalman filter/smoother-based design and implementation of digital IIR filters," *Signal Processing*, Vol. 208, 2023, pp. 108958.

[21] Stordal, A. S., Karlsen, H. A., Nævdal, G., Skaug, H. J., and Vallès, B., "Bridging the ensemble Kalman filter and particle filters: the adaptive Gaussian mixture filter," *Computational Geosciences*, Vol. 15, No. 2, 2011, pp. 293–305.

[22] Olver, F. W., Lozier, D. M., Boisvert, R. F., and Clark, C. W., "Quadrature: Gauss-Hermite formula," 2010.

[23] Naveau, P., Genton, M. G., and Shen, X., "A skewed Kalman filter," *Journal of multivariate Analysis*, Vol. 94, No. 2, 2005, pp. 382–400.

[24] Hastings, W. K., "Monte Carlo sampling methods using Markov chains and their applications," 1970.

[25] Tierney, L., "Markov chains for exploring posterior distributions," *the Annals of Statistics*, 1994, pp. 1701–1728.

[26] Bektaş, S. and Şişman, Y., "The comparison of L1 and L2-norm minimization methods," *International Journal of the Physical Sciences*, Vol. 5, No. 11, 2010, pp. 1721–1727.

[27] Hodson, T. O., "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," *Geoscientific Model Development*, Vol. 15, No. 14, 2022, pp. 5481–5487.

[28] Pastore, M. and Calcagnì, A., "Measuring distribution similarities between samples: a distribution-free overlapping index," *Frontiers in psychology*, Vol. 10, 2019, pp. 1089.

[29] Battin, R. H., "An Introduction to the Mathematics and Methods of Astrodynamics, Revised Edition," 1999.

[30] Rezaie, J. and Eidsvik, J., "A skewed unscented Kalman filter," *International Journal of Control*, Vol. 89, No. 12, 2016, pp. 2572–2583.

[31] Linacre, J. M., "Overlapping normal distributions," *Rasch Measurement Transactions*, Vol. 10, No. 1, 1996, pp. 487–8.

[32] Peralta, I. L., *PROBABILISTIC ERROR DETECTION AND DIAGNOSIS IN LARGE-SCALE*, Ph.D. thesis, Purdue University West Lafayette, 2012.

[33] Hwang, M. Y., "The E-space Inter-Domain Interaction Potential Model 3+ Sigma Confidence Level," .

[34] Pizá, Á. G., Farfán, F. D., Albarracín, A. L., Ruiz, G. A., and Felice, C. J., "Discriminability measures and time–frequency features: An application to vibrissal tactile discrimination," *Journal of Neuroscience Methods*, Vol. 233, 2014, pp. 78–88.

[35] Mulekar, M. S. and Mishra, S. N., "Overlap coefficients of two normal densities: equal means case," *Journal of the Japan Statistical Society, Japanese Issue*, Vol. 24, No. 2, 1994, pp. 169–180.

[36] Crassidis, J. L. and Markley, F. L., "Unscented filtering for spacecraft attitude estimation," *Journal of guidance, control, and dynamics*, Vol. 26, No. 4, 2003, pp. 536–542.

[37] Holmes, S., Klein, G., and Murray, D. W., "A square root unscented Kalman filter for visual monoSLAM," *2008 IEEE international conference on robotics and automation*, IEEE, 2008, pp. 3710–3716.

[38] Wu, Y., Hu, D., Wu, M., and Hu, X., "Unscented Kalman filtering for additive noise case: augmented vs. non-augmented," *Proceedings of the 2005, American Control Conference, 2005.*, IEEE, 2005, pp. 4051–4055.

APPENDICES

APPENDIX A
KALMAN FILTER ALGORITHM

At each step, the Kalman filter [10] begins by propagating the mean of the state vector $\hat{X}$ and its covariance matrix P forward in time by,

$$\hat{\mathbf{X}}(t^-) = \Phi(t, t_0)\hat{\mathbf{X}}(t_0^-), \tag{A.1}$$

$$P(t^-) = \Phi(t, t_0)P(t_0^-)\Phi^T(t, t_0) + \int_{t_0}^{t} \Phi(\tau, t_0)GQG^T\Phi^T(\tau, t_0)\,\mathrm{d}\tau, \tag{A.2}$$

where Φ is the state transition matrix of the dynamical system.

Once a measurement $\tilde{Z}$ is available, the statistics are updated by

$$\begin{aligned} K(t) &= P(t^-)H^T(t)[H(t)P(t^-)H^T(t) + R]^{-1}, \\ \hat{\mathbf{X}}(t^+) &= \hat{\mathbf{X}}(t^-) + K(t)(\tilde{\mathbf{Z}} - \hat{\mathbf{Z}}), \\ P(t^+) &= [I - K(t)H(t)]P(t^-)[I - K(t)H(t)]^T + K(t)RK^T(t), \end{aligned} \tag{A.3}$$

where H is the measurement geometry matrix, and R is the covariance matrix of the measurement noise. Derivation of the Kalman update equations (A.3) was based on both the state vector and measurement probability density functions being Gaussian. The dynamics were also assumed linear, so that if the initial distribution is Gaussian, it will remain Gaussian as it propagates forward-in-time.

A.1 EXTENDED KALMAN FILTER

In the situation where nonlinearities are present in the dynamics and measurement model, the Kalman filter can be reformulated based on first order approximations. This reformulation is called the Extended Kalman filter [10]. The first order approximation of

the propagation of the statistics through nonlinear dynamics is as follows. The nonlinear dynamics are defined by the integral equation

$$\hat{\mathbf{X}}(t^-) = \hat{\mathbf{X}}(t_0^-) + \int_{t_0}^{t} \dot{\mathbf{X}}(\mathbf{X}(\tau), \tau)\, \mathrm{d}\tau. \tag{A.4}$$

The State Transition Matrix (STM) Φ is also the Jacobian of the dynamics, and is given by,

$$\Phi(t, t_0^-) = \left.\frac{\partial \mathbf{X}(t)}{\partial \mathbf{X}(t_0)}\right|_{\mathbf{X}(t_0)=\hat{\mathbf{X}}(t_0^-)}. \tag{A.5}$$

It is calculated by integrating through time based on a formulation from the Leibniz's Integral Rule. The Leibniz's Integral Rule is given by

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_a^b f(x,t)\,\mathrm{d}t\right) = \int_a^b \frac{\partial}{\partial x} f(x,t)\,\mathrm{d}t.$$

Applying the chain rule to obtain the derivative with respect to the initial condition(s),

$$\frac{\mathrm{d}}{\mathrm{d}x_0}\left(\int_a^b f(x,t)\,\mathrm{d}t\right) = \int_a^b \left(\frac{\partial}{\partial x} f(x,t)\right)\left(\frac{\partial x}{\partial x_0}\right)\mathrm{d}t. \tag{A.6}$$

Therefore, the STM is calculated by combining (A.4), (A.5), and (A.6), to give

$$\Phi(t, t_0^-) = I + \int_{t_0}^{t} \left(\frac{\partial \dot{\mathbf{X}}}{\partial \mathbf{X}}\right) \cdot \Phi(\tau, t_0^-)\,\mathrm{d}\tau.$$

Next recall that the statistics of the state vector at t_0 are defined by

$$\begin{aligned}
\hat{\mathbf{X}}(t_0^-) &= \mathbb{E}[\mathbf{X}(t_0^-)], \\
P(t_0^-) &= \mathbb{E}\Big[(\mathbf{X}(t_0^-) - \mathbb{E}[\mathbf{X}(t_0^-)])(\mathbf{X}(t_0^-) - \mathbb{E}[\mathbf{X}(t_0^-)])^T\Big].
\end{aligned}$$

Applying a first order Taylor expansion to the dynamics centered around $\hat{X}(t_0^-)$, we find

$$\mathbf{X}(t^-) = \hat{\mathbf{X}}(t^-) + \Phi(t, t_0^-)(\mathbf{X}(t_0^-) - \hat{\mathbf{X}}(t_0^-)).$$

The propagated mean value and covariance matrix will then be calculated by using the expectation operator. First, the mean value will be determined,

$$\begin{aligned}
\mathbb{E}[\mathbf{X}(t^-)] &= \mathbb{E}[\hat{\mathbf{X}}(t^-) + \Phi(t, t_0^-)(\mathbf{X}(t_0^-) - \hat{\mathbf{X}}(t_0^-))] \\
&= \mathbb{E}[\hat{\mathbf{X}}(t^-)] + \Phi(t, t_0^-)\mathbb{E}[\mathbf{X}(t_0^-)] - \Phi(t, t_0^-)\mathbb{E}[\hat{\mathbf{X}}(t_0^-)] \\
&= \hat{\mathbf{X}}(t^-).
\end{aligned}$$

Then the covariance matrix before incorporating process noise is obtained as follows,

$$\begin{aligned}
P(t^-) &= \mathbb{E}[(\mathbf{X}(t^-) - \mathbb{E}[\mathbf{X}(t^-)])(\mathbf{X}(t^-) - \mathbb{E}[\mathbf{X}(t^-)])^T] \\
&= \mathbb{E}[\Phi(t, t_0^-)(\mathbf{X}(t_0^-) - E[\mathbf{X}(t_0^-)]) \cdot (\mathbf{X}(t_0^-) - \mathbb{E}[\mathbf{X}(t_0^-)])^T \Phi^T(t, t_0^-)] \\
&= \Phi(t, t_0^-)\mathbb{E}[(\mathbf{X}(t_0^-) - \mathbb{E}[\mathbf{X}(t_0^-)]) \cdot (\mathbf{X}(t_0^-) - \mathbb{E}[\mathbf{X}(t_0^-)])^T]\Phi^T(t, t_0^-) \\
&= \Phi(t, t_0^-)P(t_0^-)\Phi^T(t, t_0^-)
\end{aligned}$$

Adding the process noise to the covariance matrix will give the same result as (A.2).

$$P(t^-) = \Phi(t, t_0^-)P(t_0^-)\Phi^T(t, t_0^-) + \int_{t_0}^{t} \Phi(\tau, t_0^-)GQG^T\Phi^T(\tau, t_0^-)\,\mathrm{d}\tau.$$

When a measurement is available, the mean and covariance will be updated following the same linearization methodology. The nonlinear measurement model and measurement Jacobian are the following, respectively,

$$\begin{aligned}
\mathbf{Z} &= h(\mathbf{X}(t)), \\
H(t) &= \left.\frac{\partial \mathbf{Z}}{\partial \mathbf{X}(t)}\right|_{\mathbf{X}(t)=\hat{\mathbf{X}}(t)}.
\end{aligned}$$

In the same manner as the original Kalman filter, the update equations are the following,

$$
\begin{aligned}
K(t) &= P(t^-)H^T(t)[H(t)P(t^-)H^T(t) + R]^{-1}, \\
\hat{\mathbf{X}}(t^+) &= \hat{\mathbf{X}}(t^-) + K(t)(\tilde{\mathbf{Z}} - \hat{\mathbf{Z}}), \\
P(t^+) &= [I - K(t)H(t)]P(t^-)[I - K(t)H(t)]^T + K(t)RK^T(t).
\end{aligned}
$$

A.2 VERIFICATION OF KALMAN UPDATE FOUNDATION

The Kalman filter is an optimal state estimation or tracking algorithm provided that both the dynamics and measurement models are linear, mathematically, and have defined PDFs that are Gaussian [10]. When a measurement is available, the Kalman update equations are used to process the measurement and refine the state vector statistics. The foundation for the Kalman update equations is that once a measurement is received, the Gaussian distribution which had previously fully described the known statistics of the state vector must now be augmented by new information from the Gaussian distribution of the measurement model. Essentially, the probability of a region in the state space has a value obtained from the prior state PDF and a different value obtained from the measurement PDF but the true value is obtained by taking these two values and multiplying them together. This product then has to be normalized by some factor. Therefore, the new distribution will be the product of the two Gaussian distributions after being normalized by the integration of this product of the whole state space. It turns out that this updated PDF will also be a Gaussian. The following describes this mathematically, Let $h(\mathbf{X})$ be the nonlinear function for the measurement.

$$\hat{\mathbf{Z}} = h(\hat{\mathbf{X}}^-).$$

The measurement Jacobian H, the measurement estimate $\hat{\mathbf{Z}}$, and the actual measurement received $\tilde{\mathbf{Z}}$ will be used in the measurement PDF as seen by the following,

$$\phi\left(\hat{\mathbf{Z}} + H(\mathbf{X} - \hat{\mathbf{X}}^-); \tilde{\mathbf{Z}}, R\right).$$

$$\phi\left(\mathbf{X};\hat{\mathbf{X}}^{+},P^{+}\right)=\frac{\phi\left(\mathbf{X};\hat{\mathbf{X}}^{-},P^{-}\right)\cdot\phi\left(\hat{\mathbf{Z}}+H(\mathbf{X}-\hat{\mathbf{X}}^{-});\tilde{\mathbf{Z}},R\right)}{\int_{-\infty}^{\infty}\phi\left(\mathbf{X};\hat{\mathbf{X}}^{-},P^{-}\right)\cdot\phi\left(\hat{\mathbf{Z}}+H(\mathbf{X}-\hat{\mathbf{X}}^{-});\tilde{\mathbf{Z}},R\right)\mathrm{d}\mathbf{X}} \tag{A.7}$$

In order to demonstrate how the Kalman update equations represent a Gaussian distribution equivalent to the PDF in equation (A.7), first the mathematical equation for a Gaussian equation will be used to demonstrate how the update PDF has the form of a Gaussian. The following are the mathematical expressions for the two Gaussian distributions used in (A.7).

$$\begin{aligned}\phi\left(\hat{\mathbf{Z}}+H(\mathbf{X}-\hat{\mathbf{X}}^{-});\tilde{\mathbf{Z}},R\right)&=\frac{\exp\left(-\frac{1}{2}((\hat{\mathbf{Z}}-\tilde{\mathbf{Z}})+H(\mathbf{X}-\hat{\mathbf{X}}^{-}))^{T}R^{-1}((\hat{\mathbf{Z}}-\tilde{\mathbf{Z}})+H(\mathbf{X}-\hat{\mathbf{X}}^{-}))\right)}{\sqrt{\left|2\pi R\right|}}\\ \phi\left(\mathbf{X};\hat{\mathbf{X}}^{-},P^{-}\right)&=\frac{\exp\left(-\frac{1}{2}(\mathbf{X}-\hat{\mathbf{X}}^{-})^{T}(P^{-})^{-1}(\mathbf{X}-\hat{\mathbf{X}}^{-})\right)}{\sqrt{\left|2\pi P^{-}\right|}}\end{aligned} \tag{A.8}$$

The denominators in equation set (A.8) are constants and the denominator in (A.7) is also a constant. Also, the product of two exponentials can be simplified as a single exponential by the identity, $\exp(a)\exp(b)=\exp(a+b)$. Therefore, when the two equations (A.8) are substituted into (A.7), the updated PDF will have the form,

$$\phi\left(\mathbf{X};\hat{\mathbf{X}}^{+},P^{+}\right)=\eta\cdot\exp(\zeta(\mathbf{X})) \tag{A.9}$$

where η is a constant. Equation (A.9) appears to be Gaussian in form but to demonstrate that the exponent ζ can be represented as the same quadratic form as in a Gaussian, the derivative of ζ will be taken twice. The first derivative will be set equal to zero and solved for $\mathbf{X}$ to obtain the mean value and then the second derivative will be taken to find the covariance and also to show that it is a constant value which demonstrates that ζ is a quadratic equation. The following sequence of equations are each steps to find the mean

value.

$$\zeta = -\frac{1}{2}(\mathbf{X}-\hat{\mathbf{X}}^-)^T(P^-)^{-1}(\mathbf{X}-\hat{\mathbf{X}}^-) - \frac{1}{2}((\hat{\mathbf{Z}}-\tilde{\mathbf{Z}})+H(\mathbf{X}-\hat{\mathbf{X}}^-))^T R^{-1}((\hat{\mathbf{Z}}-\tilde{\mathbf{Z}})+H(\mathbf{X}-\hat{\mathbf{X}}^-))$$

$$\frac{\mathrm{d}\zeta}{\mathrm{d}\mathbf{X}} = -(\mathbf{X}-\hat{\mathbf{X}}^-)^T(P^-)^{-1} - ((\hat{\mathbf{Z}}-\tilde{\mathbf{Z}})+H(\mathbf{X}-\hat{\mathbf{X}}^-))^T R^{-1}H$$

Set $\frac{\mathrm{d}\zeta}{\mathrm{d}\mathbf{X}}$ equal to zero and solve for $\mathbf{X}$.

$$0 = (\mathbf{X}-\hat{\mathbf{X}}^-)^T(P^-)^{-1} + ((\hat{\mathbf{Z}}-\tilde{\mathbf{Z}})+H(\mathbf{X}-\hat{\mathbf{X}}^-))^T R^{-1}H$$

$$0 = (P^-)^{-T}(\mathbf{X}-\hat{\mathbf{X}}^-) + H^T R^{-T}((\hat{\mathbf{Z}}-\tilde{\mathbf{Z}})+H(\mathbf{X}-\hat{\mathbf{X}}^-))$$

$$0 = (P^-)^{-1}(\mathbf{X}-\hat{\mathbf{X}}^-) + H^T R^{-1}H(\mathbf{X}-\hat{\mathbf{X}}^-) + H^T R^{-1}(\hat{\mathbf{Z}}-\tilde{\mathbf{Z}})$$

$$((P^-)^{-1} + H^T R^{-1}H)(\mathbf{X}-\hat{\mathbf{X}}^-) = H^T R^{-1}(\tilde{\mathbf{Z}}-\hat{\mathbf{Z}})$$

$$\mathbf{X}-\hat{\mathbf{X}}^- = ((P^-)^{-1} + H^T R^{-1}H)^{-1}H^T R^{-1}(\tilde{\mathbf{Z}}-\hat{\mathbf{Z}})$$

Therefore, the updated mean value will be the following

$$\hat{\mathbf{X}}^+ = \hat{\mathbf{X}}^- + ((P^-)^{-1} + H^T R^{-1}H)^{-1}H^T R^{-1}(\tilde{\mathbf{Z}}-\hat{\mathbf{Z}}) \tag{A.10}$$

It is not immediately evident that it is equivalent to the Kalman update for the mean value, but it will be shown that Kalman Gain matrix is equal to the corresponding expression,$((P^-)^{-1}+H^T R^{-1}H)^{-1}H^T R^{-1}$, in the above equation (A.10). The following sequence of expressions will demonstrate the equivalence with the Kalman Gain matrix.

$$P^-H^T(HP^-H^T+R)^{-1} = ((P^-)^{-1} + H^T R^{-1}H)^{-1}H^T R^{-1}$$

$$P^-H^T = ((P^-)^{-1} + H^T R^{-1}H)^{-1}H^T R^{-1}(HP^-H^T+R)$$

$$P^-H^T = ((P^-)^{-1} + H^T R^{-1}H)^{-1}(H^T R^{-1}HP^-H^T + H^T)$$

$$((P^-)^{-1} + H^T R^{-1}H)P^-H^T = H^T R^{-1}HP^-H^T + H^T$$

$$H^T + H^T R^{-1}HP^-H^T = H^T R^{-1}HP^-H^T + H^T$$

Hence, the equation (A.10) is equivalent to the Kalman update equation for the mean value. Now, the covariance matrix update will be derived starting with taking the second derivative of ζ. First, the first derivative will be transposed and then differentiation will be performed. The following depicts the sequence of steps to obtain the second derivative.

$$\begin{aligned} \left(\frac{\mathrm{d}\zeta}{\mathrm{d}\mathbf{X}}\right)^T &= -(P^-)^{-1}(\mathbf{X}-\hat{\mathbf{X}}^-) - H^T R^{-1}(H(\mathbf{X}-\hat{\mathbf{X}}^-) + (\hat{Z}-\tilde{Z})) \\ \frac{\mathrm{d}}{\mathrm{d}\mathbf{X}}\left(\left(\frac{\mathrm{d}\zeta}{\mathrm{d}\mathbf{X}}\right)^T\right) &= -(P^-)^{-1} - H^T R^{-1} H \end{aligned} \tag{A.11}$$

Since the second derivative is a constant, this verifies that ζ is in fact a quadratic and thus the updated PDF (A.7) does have the form of a Gaussian distribution. It turns out that if the same steps for taking the second derivative of the exponent in a generic Gaussian distribution are followed, the covariance matrix is the inverse of the negative value of the second derivative. Therefore, following this premise, the covariance matrix for the updated PDF is the inverse of the negative value of equation (A.11), $((P^-)^{-1} + H^T R^{-1} H)^{-1}$. This will be shown to equal that of the Kalman update equations. By setting the Kalman update value equal to this expression, it will be show mathematically to be equivalent via the following steps,

$$\begin{aligned} ((P^-)^{-1} + H^T R^{-1} H)^{-1} &= (I - KH)P^- \\ ((P^-)^{-1} + H^T R^{-1} H)^{-1} &= P^- - ((P^-)^{-1} + H^T R^{-1} H)^{-1} H^T R^{-1} H P^- \\ I &= ((P^-)^{-1} + H^T R^{-1} H)P^- - H^T R^{-1} H P^- \\ I &= I + H^T R^{-1} H P^- - H^T R^{-1} H P^- \end{aligned}$$

Therefore, the covariance update derived in this section exactly matches that of the Kalman update. The mean value and covariance matrix were determined based on the understanding that a new probability density function after a measurement update would constitute a normalized product of the previously tracked Gaussian distribution of the state and the Gaussian distribution acquired from the measurement received. Since these results match

exactly with what the Kalman update claims, it validates that the Kalman filter measurement update is based on this foundation.

A.3 VERIFICATION OF JOSEPH FORM

There are two forms for the covariance update in the Kalman filter,

$$\begin{aligned} P^+ &= (I - KH)P^- \\ P^+ &= (I - KH)P^-(I - KH)^T + KRK^T \end{aligned} \tag{A.12}$$

The second equation in (A.12) is referred to as the Joseph form. The Joseph form is the preferred equation to implement into a Kalman filter algorithm because it "squares-up" the covariance matrix due to the multiplication of matrices and their respective transpose matrices. The reason this is beneficial is that during propagation of the covariance matrix through dynamical evolution, the off-diagonal term may differ than their respective transpose counterparts. The Joseph form resets the covariance matrix back to being a symmetric matrix. The purpose of this appendix is to mathematically demonstrate that the two equations in (A.12) are identical. The following sequence of steps will begin with setting the two equations equal to each other and working both sides and cancelling terms until an expression that will be equal to the zero matrix is reached at which point the known Kalman Gain matrix is inserted into the expression and worked until both sides of the equation are equal to the zero matrix.

$$\begin{aligned} (I - KH)P^- &= (I - KH)P^-(I - KH)^T + KRK^T \\ P^- - KHP^- &= (P^- - KHP^-)(I - KH)^T + KRK^T \\ P^- - KHP^- &= P^- - P^-H^TK^T - KHP^- + KHP^-H^TK^T + KRK^T \\ 0 &= -P^-H^TK^T + KHP^-H^TK^T + KRK^T \end{aligned} \tag{A.13}$$

Now inserting the known expression for the Kalman gain to modify the term KRK^T,

$$\begin{aligned} K &= P^- H^T (HP^- H^T + R)^{-1} \\ K(HP^- H^T + R) &= P^- H^T \\ KHP^- H^T + KR &= P^- H^T \\ KR &= P^- H^T - KHP^- H^T \\ KRK^T &= P^- H^T K^T - KHP^- H^T K^T \end{aligned} \tag{A.14}$$

Now, inserting (A.14) into (A.13),

$$-P^- H^T K^T + KHP^- H^T K^T + P^- H^T K^T - KHP^- H^T K^T = 0 \tag{A.15}$$

Since both sides of (A.15) are equal to the zero matrix, this means that the Joseph form for the covariance update in the Kalman update has been proven to be valid.

APPENDIX B

STEP-BACK KALMAN FILTER EQUIVALENCE FOR LINEAR SYSTEMS

The premise for the SBKF is to apply the measurement update at a time in the past where the probability distribution of the state is Gaussian, or at least well-approximated by a Gaussian. The distribution at the start of all the previously described filter algorithms is modeled as a Gaussian. After a measurement update is performed, the resulting mean and covariance are then modeled as describing a new Gaussian distribution. The SBKF begins exactly the same as the Kalman filter by propagating forward in time and obtaining a measurement prediction, $\hat{Z}$, and measurement Jacobian, $H(t)$. Using the chain rule, the measurement Jacobian with respect to the state vector at time t_0 can be calculated by

$$H(t_0) = H(t)\Phi(t, t_0). \tag{B.1}$$

Also, the contribution of process noise added to the covariance matrix during propagation from t_0 to t is given as,

$$P_q = \int_{t_0}^{t} \Phi(\tau, t_0) G Q G^T \Phi^T(\tau, t_0)\, \mathrm{d}\tau.$$

Applying $H(t_0)$ and P_q to find the Kalman update at time t_0 and corresponding statistics at time t is given by

$$\begin{aligned}
P_{q0} &= \Phi^{-1}(t,t_0)P_q\Phi^{-T}(t,t_0),\\
K(t_0) &= [P(t_0^-)+P_{q0}]H^T(t_0)\cdot[H(t_0)[P(t_0^-)+P_{q0}]H^T(t_0)+R]^{-1},\\
\hat{X}(t_0^+) &= \hat{X}(t_0^-)+K(t_0)(\tilde{Z}-\hat{Z}),\\
\hat{X}(t^+) &= \Phi(t,t_0)\hat{X}(t_0^+),\\
P(t_0^+) &= [I-K(t_0)H(t_0)][P(t_0^-)+P_{q0}][I-K(t_0)H(t_0)]^T+K(t_0)RK^T(t_0),\\
P(t^+) &= \Phi(t,t_0)P(t_0^+)\Phi^T(t,t_0).
\end{aligned} \tag{B.2}$$

In the case where the both the dynamics and measurement models are linear, the SBKF can be shown to yield identical results as the Kalman filter for $\hat{X}(t^+)$ and $P(t^+)$. Equation set (B.2) will match equation set (3.17) in a linear system as demonstrated by the following arguments.

Inserting the measurement Jacobian relationship (B.1) into the Kalman Gain at t_0,

$$K(t_0) = [P(t_0^-)+P_{q0}]\Phi^T(t,t_0)H^T(t)[H(t)\Phi(t,t_0)[P(t_0^-)+P_{q0}]\Phi^T(t,t_0)H^T(t)+R]^{-1}.$$

With some simplifcation, we find

$$\begin{aligned}
&\Phi(t,t_0)[P(t_0^-)+P_{q0}]\Phi^T(t,t_0)\\
&\qquad = \Phi(t,t_0)P(t_0^-)\Phi^T(t,t_0)+\Phi(t,t_0)P_{q0}\Phi^T(t,t_0)\\
&\qquad = P(t^-).
\end{aligned} \tag{B.3}$$

Therefore, if $K(t_0)$ is multiplied by $\Phi(t,t_0)$, then

$$\Phi(t,t_0)K(t_0) = K(t). \tag{B.4}$$

Applying (B.4) toward $\hat{X}(t^+)$ as given in (B.2),

$$\begin{aligned}\hat{X}(t^+) &= \Phi(t,t_0)\hat{X}(t_0^+) \\ &= \Phi(t,t_0)[\hat{X}(t_0^-) + K(t_0)(\tilde{Z} - \hat{Z})] \\ &= \hat{X}(t^-) + K(t)(\tilde{Z} - \hat{Z}).\end{aligned}$$

This matches $\hat{X}(t^+)$ as given in (3.17). Now, the covariance matrix $P(t^+)$ will be demonstrated to match by similar arguments. We modify expressions as follows

$$\begin{aligned}&\Phi(t,t_0)[I - K(t_0)H(t_0)] \\ &\qquad = \Phi(t,t_0)[I - K(t_0)H(t)\Phi(t,t_0)] \\ &\qquad = \Phi(t,t_0) - K(t)H(t)\Phi(t,t_0) \\ &\qquad = [I - K(t)H(t)]\Phi(t,t_0),\end{aligned} \tag{B.5}$$

By expanding $P(t^+)$ represented in (B.2) and applying (B.3), (B.4), and (B.5),

$$\begin{aligned}&P(t^+) \\ &= \Phi(t,t_0)\Big[[I - K(t_0)H(t_0)][P(t_0^-) + P_{q0}][I - K(t_0)H(t_0)]^T + K(t_0)RK^T(t_0)\Big]\Phi^T(t,t_0) \\ &= [I - K(t)H(t)]\Phi(t,t_0)[P(t_0^-) + P_{q0}]\Phi^T(t,t_0)[I - K(t)H(t)]^T + \Phi(t,t_0)K(t_0)RK^T(t_0)\Phi^T(t,t_0) \\ &= [I - K(t)H(t)]P(t^-)[I - K(t)H(t)]^T + K(t)RK^T(t).\end{aligned} \tag{B.6}$$

Therefore, $P(t^+)$ in (B.6) matches the same expression in (3.17). The equivalence between the Kalman filter and the Step-Back Kalman filter in a linear system has now been demonstrated, however, the primary purpose of development of the SBKF is the significant advantage over the KF when using the extended forms, EKF and ESBKF. The extended forms are used when nonlinear dynamics and/or nonlinear measurements are present.

APPENDIX C

DISCRETE PROCESS NOISE COVARIANCE MATRIX

Process noise inherent in the dynamics is handled in Extended Kalman filter and Extended Step-Back Kalman filter by including an extra term in the differential equation, GQG^T, used to propagate the covariance matrix,

$$P(t^+) = P(t_0^+) + \int_{t_0}^{t} FP + PF^T + GQG^T d\tau.$$

The discrete form of this propagation is the following,

$$P(t^+) = \phi P(t_0^+)\phi^T + \int_{t_0}^{t} \phi GQG^T \phi^T d\tau.$$

Hence, the discrete process noise covariance matrix is the term

$$\int_{t_0}^{t} \phi GQG^T \phi^T d\tau. \tag{C.1}$$

These equations for propagating the covariance matrix whether or not process noise is included are based on a linear dynamical system. Hence, for the orbital dynamics model, the discrete and continuous form will differ somewhat. The Unscented Kalman filter uses sigma points to gather sample statistics. Since the process noise is not included when propagating each sigma point, the process noise contribution is incorporated via a discrete covariance matrix following (C.1). Based on the algorithm for the UKF [38], each sigma-point is propagated forward one time step without process noise in the dynamics. After the propagation of each sigma-point, the sample mean and covariance matrix are calculated. A discrete covariance matrix term representing the process noise contribution is added to the sample covariance matrix. Then, using this revised sample covariance matrix along with the sample mean, new sigma-points are acquired and are then either propagated to the next time

step or, if a measurement is available, used to update the statistics from a measurement. The goal is to develop the discrete process noise covariance matrix. This can be acquired by using a second-order approximation for the dynamics propagation, obtaining a formula for the state transition matrix, and then integrating according to equation (C.1). By applying Heun's method, a second-order approximation for the dynamics can be obtained, as per the following.

The magnitude of the position vector at the previous time step is,

$$r = \|\mathbf{r}\|$$

After applying Heun's method, the dynamical model becomes,

$$\begin{aligned} \mathbf{r}^* &= \mathbf{r} + \mathbf{v}\Delta t - \frac{\mu \Delta t^2}{2}\left(\frac{\mathbf{r}}{r^3}\right) \\ \mathbf{v}^* &= \mathbf{v} - \frac{\mu \Delta t}{2}\left(\frac{\mathbf{r}}{r^3} + \frac{\mathbf{r} + \mathbf{v}\Delta t}{\|\mathbf{r} + \mathbf{v}\Delta t\|^3}\right) \end{aligned} \tag{C.2}$$

In order to find the state transition matrix, we must first find the formula for $\frac{\partial}{\partial \mathbf{r}}\left(\frac{\mathbf{r}}{r^3}\right)$.

$$\begin{aligned} \frac{\partial}{\partial \mathbf{r}}\left(\frac{\mathbf{r}}{r^3}\right) &= \frac{1}{r^3} I_{3\times 3} + \mathbf{r}\frac{\partial}{\partial \mathbf{r}}\left(\frac{1}{r^3}\right) \\ &= \frac{1}{r^3} I_{3\times 3} + \mathbf{r}\frac{\partial}{\partial r}\left(\frac{1}{r^3}\right)\frac{\partial r}{\partial \mathbf{r}} \\ &= \frac{1}{r^3} I_{3\times 3} + \mathbf{r}\left(-3r^{-4}\hat{\mathbf{i}}_r^T\right) \\ &= \frac{1}{r^3}\left(I_{3\times 3} - 3\hat{\mathbf{i}}_r\hat{\mathbf{i}}_r^T\right) \end{aligned}$$

Now, obtaining the four Jacobian matrices that make up the state transition matrix,

$$\begin{aligned} \frac{\partial \mathbf{r}^*}{\partial \mathbf{r}} &= I_{3\times 3} - \frac{\mu \Delta t^2}{2}\frac{1}{r^3}\left(I_{3\times 3} - 3\hat{\mathbf{i}}_r\hat{\mathbf{i}}_r^T\right) \\ \frac{\partial \mathbf{r}^*}{\partial \mathbf{v}} &= \Delta t\ I_{3\times 3} \end{aligned}$$

let, $\mathbf{r}_1 = \mathbf{r} + \mathbf{v}\Delta t$

$$\frac{\partial \mathbf{v}^*}{\partial \mathbf{r}} = -\frac{\mu \Delta t}{2}\frac{1}{r^3}\left(I_{3\times 3} - 3\hat{\mathbf{i}}_r \hat{\mathbf{i}}_r^T\right) - \frac{\mu \Delta t}{2}\frac{1}{r_1^3}\left(I_{3\times 3} - 3\hat{\mathbf{i}}_{r_1} \hat{\mathbf{i}}_{r_1}^T\right)$$
$$\frac{\partial \mathbf{v}^*}{\partial \mathbf{v}} = I_{3\times 3} - \frac{\mu \Delta t^2}{2}\frac{1}{r_1^3}\left(I_{3\times 3} - 3\hat{\mathbf{i}}_{r_1} \hat{\mathbf{i}}_{r_1}^T\right)$$

Combining these matrices to create the state transition matrix yields,

$$\phi = \begin{bmatrix} \frac{\partial \mathbf{r}^*}{\partial \mathbf{r}} & \frac{\partial \mathbf{r}^*}{\partial \mathbf{v}} \\ \frac{\partial \mathbf{v}^*}{\partial \mathbf{r}} & \frac{\partial \mathbf{v}^*}{\partial \mathbf{v}} \end{bmatrix}$$

. Finding the integrand in (C.1),

$$GQG^T = \begin{bmatrix} 0_{3\times 3} & 0_{3\times 3} \\ 0_{3\times 3} & Q \end{bmatrix}$$
$$\phi GQG^T\phi^T = \begin{bmatrix} \left(\frac{\partial \mathbf{r}^*}{\partial \mathbf{v}}\right) Q \left(\frac{\partial \mathbf{r}^*}{\partial \mathbf{v}}\right)^T & \left(\frac{\partial \mathbf{r}^*}{\partial \mathbf{v}}\right) Q \left(\frac{\partial \mathbf{v}^*}{\partial \mathbf{v}}\right)^T \\ \left(\frac{\partial \mathbf{v}^*}{\partial \mathbf{v}}\right) Q \left(\frac{\partial \mathbf{r}^*}{\partial \mathbf{v}}\right)^T & \left(\frac{\partial \mathbf{v}^*}{\partial \mathbf{v}}\right) Q \left(\frac{\partial \mathbf{v}^*}{\partial \mathbf{v}}\right)^T \end{bmatrix}$$

To simplify further, from numerical analysis the following holds by several orders of magnitude,

$$\left|-\frac{\mu \Delta t^2}{2}\frac{1}{r^3}\left(I_{3\times 3} - 3\hat{\mathbf{i}}_r \hat{\mathbf{i}}_r^T\right)\right| << |I_{3\times 3}|$$

$$\left|-\frac{\mu \Delta t^2}{2}\frac{1}{r_1^3}\left(I_{3\times 3} - 3\hat{\mathbf{i}}_{r_1} \hat{\mathbf{i}}_{r_1}^T\right)\right| << |I_{3\times 3}|$$

Therefore,

$$\phi GQG^T\phi^T = \begin{bmatrix} Q\Delta t^2 & Q\Delta t \\ Q\Delta t & Q \end{bmatrix}$$

Finally, evaluating the integral in (C.1),

$$\int_0^{\Delta t} \phi G Q G^T \phi^T d\tau = \begin{bmatrix} Q\frac{\Delta t^3}{3} & Q\frac{\Delta t^2}{2} \\ Q\frac{\Delta t^2}{2} & Q\Delta t \end{bmatrix}. \tag{C.3}$$

C.1 NUMERICAL VERIFICATION

To verify that the explicit approximation for the discrete process noise covariance matrix is accurate over a reasonable time step, a Monte Carlo simulation running actual orbital dynamics with process noise in the acceleration was implemented and compared against (C.3). The parameters of the simulation are the following,

$$\mathbf{r}_0 = \begin{bmatrix} 39621.502751988 & 14420.7554640182 & 63.8935010377406 \end{bmatrix}^T \boldsymbol{km}$$

$$\mathbf{v}_0 = \begin{bmatrix} -1.05154888245646 & 2.88911713863564 & 0.0264234241457446 \end{bmatrix}^T \frac{\boldsymbol{km}}{\boldsymbol{s}}$$

$$Q = 10^{-13} \begin{bmatrix} 0.20 & 0.00 & 0.00 \\ 0.00 & 0.30 & 0.00 \\ 0.00 & 0.00 & 0.25 \end{bmatrix} \frac{\boldsymbol{km^2}}{\boldsymbol{s^4}}$$

$$\Delta t = 300 \ \boldsymbol{s}$$

$$\mu = 398600.436 \ \frac{\boldsymbol{km^3}}{\boldsymbol{s^2}}$$

The process noise covariance obtained from the Monte Carlo simulation, which generated and used one million samples, is the following,

$$10^{-6}\begin{bmatrix} 0.1800421519 & 0.0001030185 & 0.0002094364 & 0.0009003783 & -0.0000006389 & 0.000001146 \\ 0.0001030185 & 0.2698592091 & 0.0001236954 & 0.0000013574 & 0.0013497446 & 0.000001375 \\ 0.0002094364 & 0.0001236954 & 0.2248234685 & 0.0000010769 & 0.0000010276 & 0.001123535 \\ 0.0009003783 & 0.0000013574 & 0.0000010769 & 0.0000059989 & -0.0000000012 & 0.000000005 \\ -0.0000006389 & 0.0013497446 & 0.0000010276 & -0.0000000012 & 0.0000089958 & 0.000000008 \\ 0.0000011465 & 0.0000013758 & 0.0011235357 & 0.0000000059 & 0.0000000081 & 0.000007487 \end{bmatrix}$$

The process noise covariance from (C.3) is the following,

$$10^{-6}\begin{bmatrix} 0.1800000000 & 0.0000000000 & 0.0000000000 & 0.0009000000 & 0.0000000000 & 0.0000000000 \\ 0.0000000000 & 0.2700000000 & 0.0000000000 & 0.0000000000 & 0.0013500000 & 0.0000000000 \\ 0.0000000000 & 0.0000000000 & 0.2250000000 & 0.0000000000 & 0.0000000000 & 0.0011250000 \\ 0.0009000000 & 0.0000000000 & 0.0000000000 & 0.0000060000 & 0.0000000000 & 0.0000000000 \\ 0.0000000000 & 0.0013500000 & 0.0000000000 & 0.0000000000 & 0.0000090000 & 0.0000000000 \\ 0.0000000000 & 0.0000000000 & 0.0011250000 & 0.0000000000 & 0.0000000000 & 0.0000075000 \end{bmatrix}$$

When comparing the two matrices, they closely match therefore confirming that (C.3) is accurate.

APPENDIX D

MEAN AND COVARIANCE IN GAUSSIAN MIXTURE MODEL

In a Gaussian Mixture Model, the probability density function of the state is approximated as a sum of unique normal distributions. Each normal distribution is defined by a mean value, covariance matrix, and a weighting factor, however, the GMM requires evaluation of the overall mean value and covariance of the sum of the weighted normal distributions. The following will demonstrate how this is achieved. Let the GMM PDF be defined as the following,

$$p(\mathbf{X}) = \sum_{i=1}^{N} \omega_i \cdot \phi\left(\mathbf{X}; \boldsymbol{\mu}_i, P_i\right). \tag{D.1}$$

The weights ω_i in (D.1) must sum up to 1,

$$\sum_{i=1}^{N} \omega_i = 1. \tag{D.2}$$

Now, using the expected value function to find the mean value of (D.1),

$$\begin{aligned} \mathbb{E}[\mathbf{X}] &= \int_{-\infty}^{\infty} \mathbf{X} \cdot \sum_{i=1}^{N} \omega_i \cdot \phi\left(\mathbf{X}; \boldsymbol{\mu}_i, P_i\right) \mathrm{d}\mathbf{X} \\ &= \sum_{i=1}^{N} \omega_i \cdot \int_{-\infty}^{\infty} \mathbf{X} \cdot \phi\left(\mathbf{X}; \boldsymbol{\mu}_i, P_i\right) \mathrm{d}\mathbf{X} \\ &= \sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i. \end{aligned} \tag{D.3}$$

The covariance matrix will now be derived using the result of the mean value in (D.3) and using an alternative form of the covariance matrix when using the expected value function.

$$\mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T] = \mathbb{E}[\mathbf{X}\mathbf{X}^T] - \mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{X}]^T$$

Inserting value for mean value(D.3) as well as using (D.2).

$$= \left(\sum_{i=1}^{N} \omega_i\right) \mathbb{E}[\mathbf{X}\mathbf{X}^T] - \left(\sum_{i=1}^{N} \omega_i\right) \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right) \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right)^T$$

Add and subtract $\sum_{i=1}^{N} \left(\omega_i \cdot \boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right)$

$$\begin{aligned}
&= \left(\sum_{i=1}^{N} \omega_i\right) \mathbb{E}[\mathbf{X}\mathbf{X}^T] - \sum_{i=1}^{N} \left(\omega_i \cdot \boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right) \\
&+ \sum_{i=1}^{N} \left(\omega_i \cdot \boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right) - \left(\sum_{i=1}^{N} \omega_i\right) \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right) \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right)^T \\
&= \sum_{i=1}^{N} \omega_i \left(\mathbb{E}[\mathbf{X}\mathbf{X}^T] - \boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right) + \sum_{i=1}^{N} \omega_i \cdot \left(\left(\boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right) - \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right) \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right)^T\right)
\end{aligned}$$

Reworking a portion of the last term,

$$\left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right) \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right)^T = \sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i \mathbb{E}[\mathbf{X}]^T = \sum_{i=1}^{N} \omega_i \cdot \mathbb{E}[\mathbf{X}] \boldsymbol{\mu}_i{}^T$$

Now, applying this expression into the covariance formula

$$\begin{aligned}
&\sum_{i=1}^{N} \omega_i \left(\mathbb{E}[\mathbf{X}\mathbf{X}^T] - \boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right) + \sum_{i=1}^{N} \omega_i \cdot \left(\left(\boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right) - \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right) \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right)^T\right) \\
&= \sum_{i=1}^{N} \omega_i \left(\mathbb{E}[\mathbf{X}\mathbf{X}^T] - \boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right) \\
&+ \sum_{i=1}^{N} \omega_i \cdot \left(\left(\boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right) - \sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i \mathbb{E}[\mathbf{X}]^T - \sum_{i=1}^{N} \omega_i \cdot \mathbb{E}[\mathbf{X}] \boldsymbol{\mu}_i{}^T + \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right) \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i\right)^T\right)
\end{aligned}$$

The first term can be simplified as follows,

$$\sum_{i=1}^{N} \omega_i \left(\mathbb{E}[\mathbf{X}\mathbf{X}^T] - \boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T\right) = \sum_{i=1}^{N} \omega_i \cdot P_i$$

Also, in the second term, the following expressions be simplified as follows,

$$\sum_{i=1}^{N} \omega_i \cdot \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i \mathbb{E}[\mathbf{X}]^T \right) = \sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i \mathbb{E}[\mathbf{X}]^T$$
$$\sum_{i=1}^{N} \omega_i \cdot \left(\sum_{i=1}^{N} \omega_i \cdot \mathbb{E}[\mathbf{X}] \boldsymbol{\mu}_i{}^T \right) = \sum_{i=1}^{N} \omega_i \cdot \mathbb{E}[\mathbf{X}] \boldsymbol{\mu}_i{}^T$$

Applying these revised expressions into the covariance formula,

$$\sum_{i=1}^{N} \omega_i \left(\mathbb{E}[\mathbf{X}\mathbf{X}^T] - \boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T \right)$$
$$+ \sum_{i=1}^{N} \omega_i \cdot \left(\left(\boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T \right) - \sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i \mathbb{E}[\mathbf{X}]^T - \sum_{i=1}^{N} \omega_i \cdot \mathbb{E}[\mathbf{X}] \boldsymbol{\mu}_i{}^T + \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i \right) \left(\sum_{i=1}^{N} \omega_i \cdot \boldsymbol{\mu}_i \right)^T \right)$$
$$= \sum_{i=1}^{N} \omega_i \cdot P_i + \sum_{i=1}^{N} \omega_i \cdot \left(\left(\boldsymbol{\mu}_i \boldsymbol{\mu}_i{}^T \right) - \boldsymbol{\mu}_i \mathbb{E}[\mathbf{X}]^T - \mathbb{E}[\mathbf{X}] \boldsymbol{\mu}_i{}^T + \mathbb{E}[\mathbf{X}] \mathbb{E}[\mathbf{X}]^T \right)$$

Therefore, the covariance for the GMM is the following,

$$\mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}]) (\mathbf{X} - \mathbb{E}[\mathbf{X}])^T] = \sum_{i=1}^{N} \omega_i \cdot P_i + \sum_{i=1}^{N} \omega_i \cdot (\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}]) (\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T$$

An alternative approach for deriving the covariance of the GMM will now be presented. The expected value expression for the covariance matrix will be reformulated by adding and subtracting the term $\boldsymbol{\mu}_i$ in two places as follows,

$$\mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}]) (\mathbf{X} - \mathbb{E}[\mathbf{X}])^T] = \mathbb{E}[(\mathbf{X} - \boldsymbol{\mu}_i + \boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}]) (\mathbf{X} - \boldsymbol{\mu}_i + \boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T]$$

This will now be expanded,

$$= \mathbb{E}[(\mathbf{X} - \boldsymbol{\mu}_i)(\mathbf{X} - \boldsymbol{\mu}_i)^T] + \mathbb{E}[(\mathbf{X} - \boldsymbol{\mu}_i)(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T] +$$
$$\mathbb{E}[(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \boldsymbol{\mu}_i)^T] + \mathbb{E}[(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T]$$

The above expression does not specify which particle mean value, $\boldsymbol{\mu}_i$, is being applied and this was intentionally generalized. Let the function $g(\mathbf{X})$ be defined as the following.

$$g(\mathbf{X}) = \int_{\infty}^{\infty} \mathbf{X} \cdot \phi\left(\mathbf{X}; \boldsymbol{\mu}_i, P_i\right) \mathrm{d}\mathbf{X}$$

Then the covariance, based on (D.1) as the underlying PDF, can be reformulated as the following,

$$\mathbb{E}[\left(\mathbf{X} - \mathbb{E}[\mathbf{X}]\right)\left(\mathbf{X} - \mathbb{E}[\mathbf{X}]\right)^T] = \sum_{i=1}^{N} \omega_i \cdot g((\mathbf{X} - \mathbb{E}[\mathbf{X}])\left(\mathbf{X} - \mathbb{E}[\mathbf{X}]\right)^{\mathbf{T}})$$

$$= \sum_{i=1}^{N} \omega_i \cdot g((\mathbf{X} - \boldsymbol{\mu}_i)(\mathbf{X} - \boldsymbol{\mu}_i)^T) + \sum_{i=1}^{N} \omega_i \cdot g((\mathbf{X} - \boldsymbol{\mu}_i)(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T) +$$

$$\sum_{i=1}^{N} \omega_i \cdot g((\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \boldsymbol{\mu}_i)^T) + \sum_{i=1}^{N} \omega_i \cdot g((\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T)$$

If the particle corresponds with the same index as the normal distribution when the expected value function is being performed then the first and last terms are identical to the following,

$$\sum_{i=1}^{N} \omega_i \cdot g((\mathbf{X} - \boldsymbol{\mu}_i)(\mathbf{X} - \boldsymbol{\mu}_i)^T) = \sum_{i=1}^{N} \omega_i \cdot P_i$$

$$\sum_{i=1}^{N} \omega_i \cdot g((\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T) = \sum_{i=1}^{N} \omega_i \cdot (\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T$$

The middle two terms will become the following,

$$\sum_{i=1}^{N} \omega_i \cdot g((\mathbf{X} - \boldsymbol{\mu}_i)(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T) = \sum_{i=1}^{N} \omega_i \cdot g((\mathbf{X} - \boldsymbol{\mu}_i)) \cdot (\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T = 0$$

$$\sum_{i=1}^{N} \omega_i \cdot g((\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \boldsymbol{\mu}_i)^T) = \sum_{i=1}^{N} \omega_i \cdot (\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}]) \cdot g((\mathbf{X} - \boldsymbol{\mu}_i)^T) = 0$$

Using this alternative approach the covariance matrix for the GMM is once again the following,

$$\mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^T] = \sum_{i=1}^{N} \omega_i \cdot P_i + \sum_{i=1}^{N} \omega_i \cdot (\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])(\boldsymbol{\mu}_i - \mathbb{E}[\mathbf{X}])^T$$

APPENDIX E

SUFFICIENCY OF KEPLERIAN DYNAMICS

In the PROBLEM FORMULATION AND SETUP chapter, in was presented that the dynamics model used in all of the filtering algorithms is Keplerian orbital dynamics. A higher fidelity dynamics model was implemented in all the filters, however, it was computationally burdensome and when compared with the lower fidelity Keplerian model, they were practically indistinguishable from one another. This prompted the use of applying Keplerian dynamics to filters and the higher fidelity model was applied solely to the truth simulation of the RSO and OBSs. The results leading to this conclusion are depicted in this section. Table E.1 details the increase in computational requirements needed to have a higher fidelity dynamical model. The scale of the increase in computation time is significant which supported the conclusion that Keplerian dynamics would the model to implement.

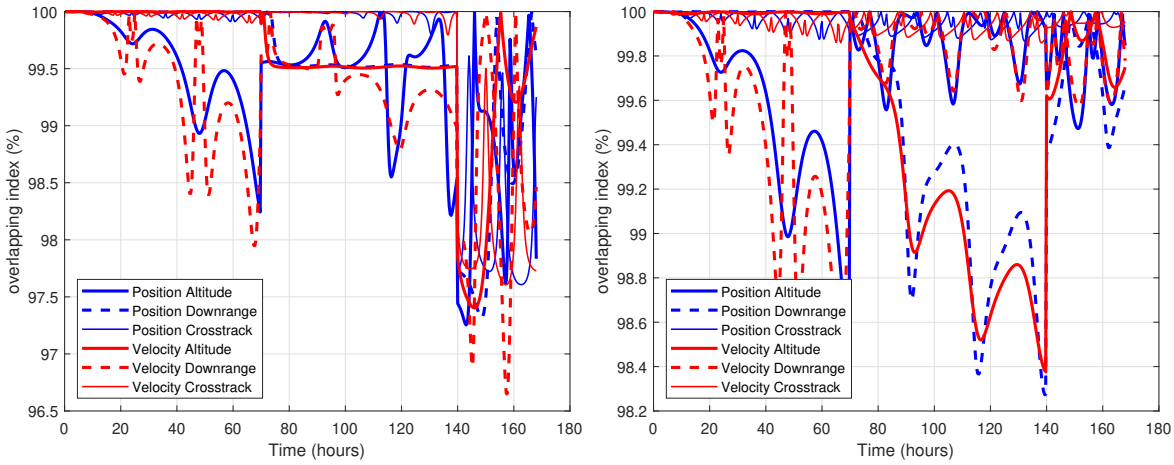
Table E.1: Computation Time Advantage of using Keplerian Dynamics

Filter Algorithm	Ratio of Full Dynamics to Keplerian Dynamics
EKF	17.28
HKF	5.70
UKF	19.15
GMM (2 nodes per dimension)	33.99
ESBKF	24.60

The following plots in this section will show η calculations which compares σ values obtained from using the higher fidelity dynamical model vs. the Keplerian dynamical model. There are two σ values, described in the Overview section of this chapter, for each dynamical model simulation. Each plot compares both σ values.

To begin, Figure E.1 shows the results of the EKF when applying full dynamics and

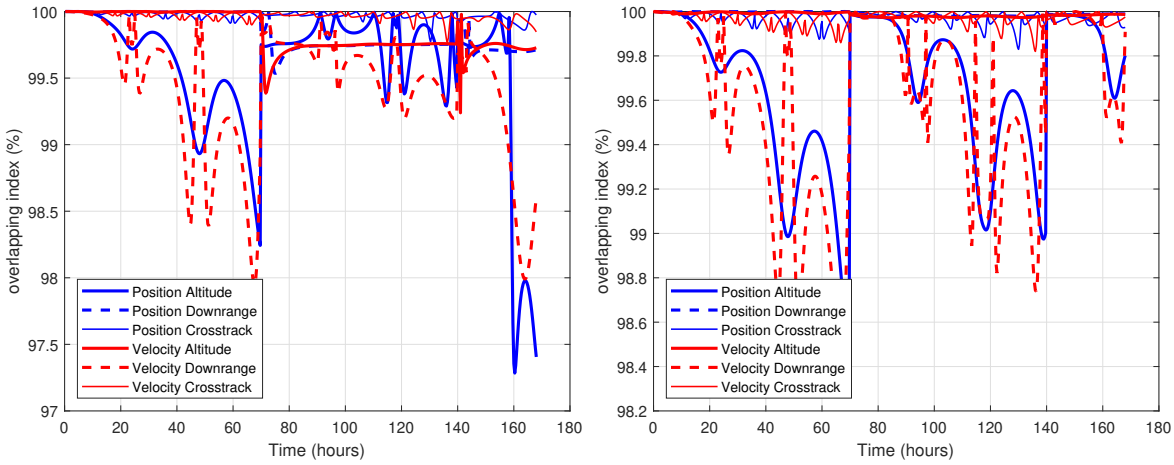
Keplerian dynamics. The η values for both σ values are in the upper 90 percentile range which concludes these figures are indistinguishable from one another.



(a) monte-carlo overlap (b) filter overlap

Fig. E.1: EKF, Full dynamics vs. Keplerian dynamics Overlap

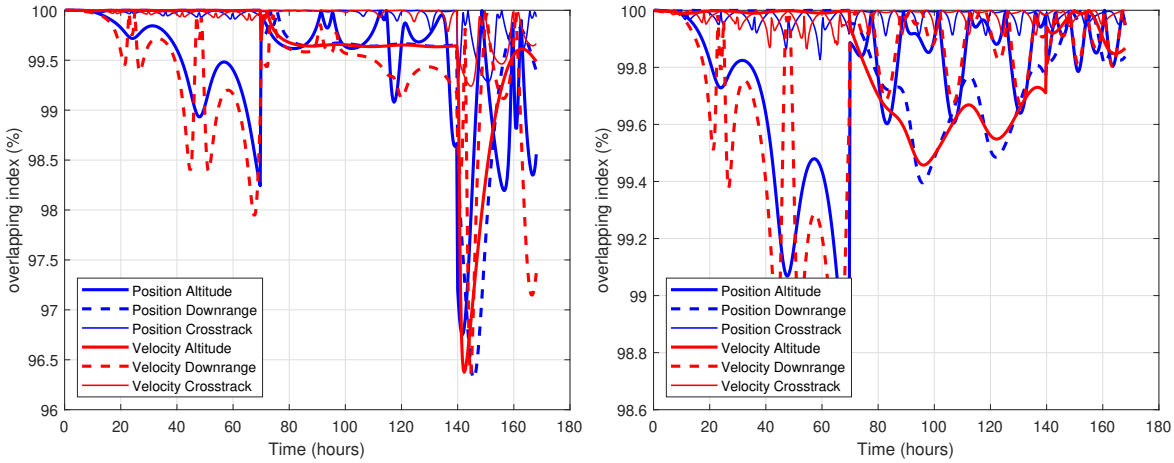
Next, the dynamical model comparison for the HKF is shown in Figure E.2 which likewise demonstrates no significant differences when using Keplerian dynamics. The η values remain in the higher 90-percentile range.



(a) monte-carlo overlap (b) filter overlap

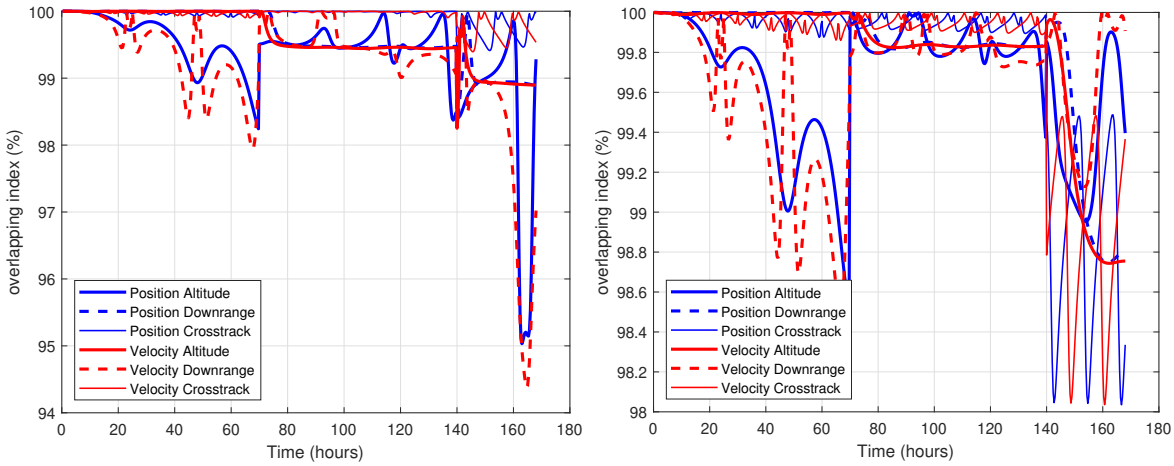
Fig. E.2: HKF, Full dynamics vs. Keplerian dynamics Overlap

Following suit with the EKF and HKF, the UKF, GMM, and ESBKF when using either Keplerian dynamics or a higher fidelity model that has additional gravitational, solar, and lunar perturbing accelerations, have no distinguishable difference in filter performance. These results are seen in Figure E.3 for the UKF, Figure E.4 for the GMM, and Figure E.5 for the ESBKF.



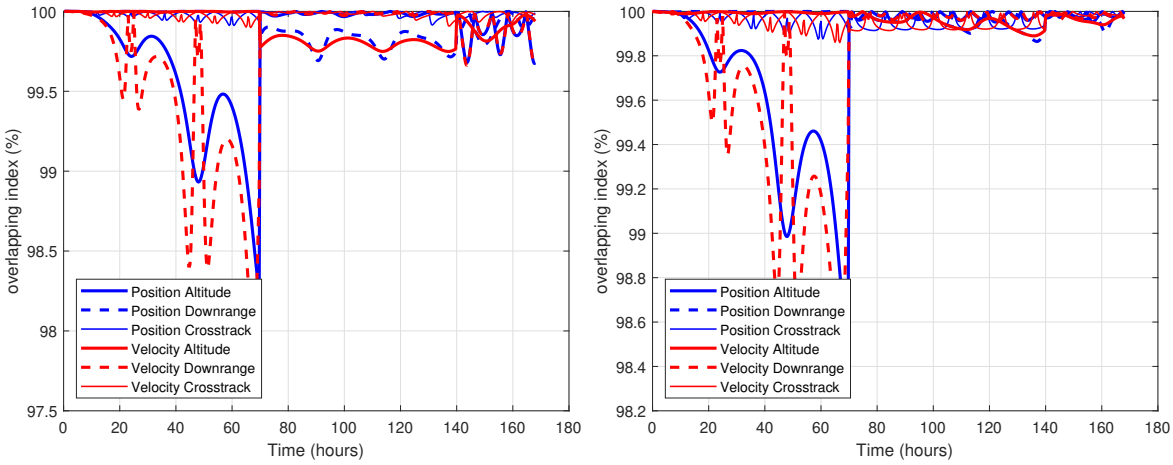
(a) monte-carlo overlap (b) filter overlap

Fig. E.3: UKF, Full dynamics vs. Keplerian dynamics Overlap



(a) monte-carlo overlap (b) filter overlap

Fig. E.4: GMM, Full dynamics vs. Keplerian dynamics Overlap



(a) monte-carlo overlap (b) filter overlap

Fig. E.5: ESBKF, Full dynamics vs. Keplerian dynamics Overlap

The conclusion for all the filters from η analysis, seen in Figures E.1-E.5, is that there is no need to add computational burden for propagating a more accurate dynamical model when a Kepler orbit produces effectively the same result and with a fraction of the computational requirements.

APPENDIX F

SUPPLEMENTAL PLOTS

F.1 EXTENDED KALMAN FILTER

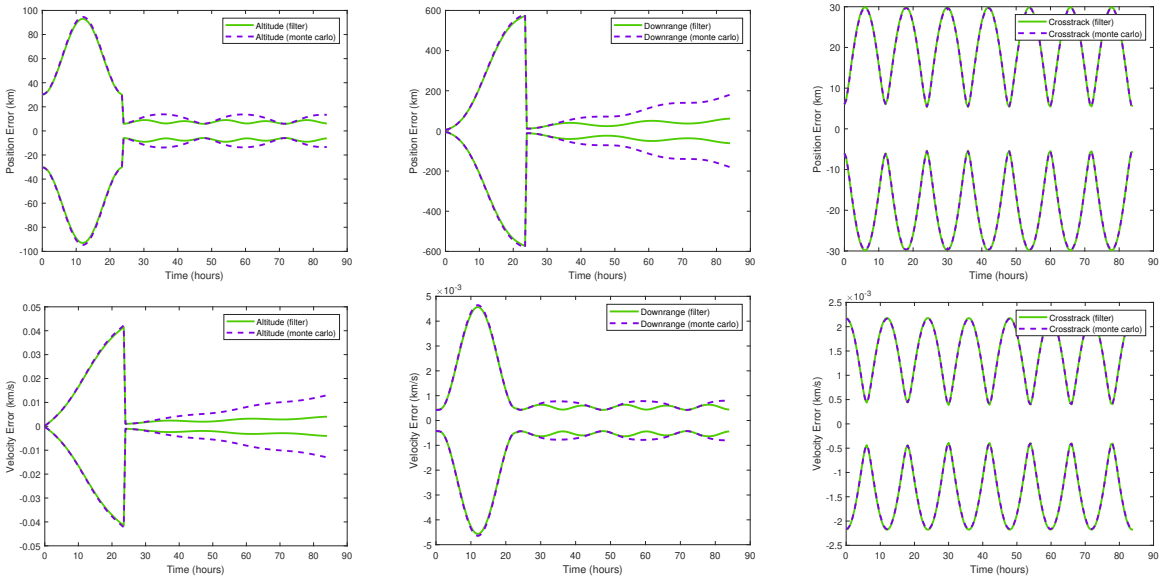


Fig. F.1: EKF Measurement at 24 Hours, OBS 1 Only

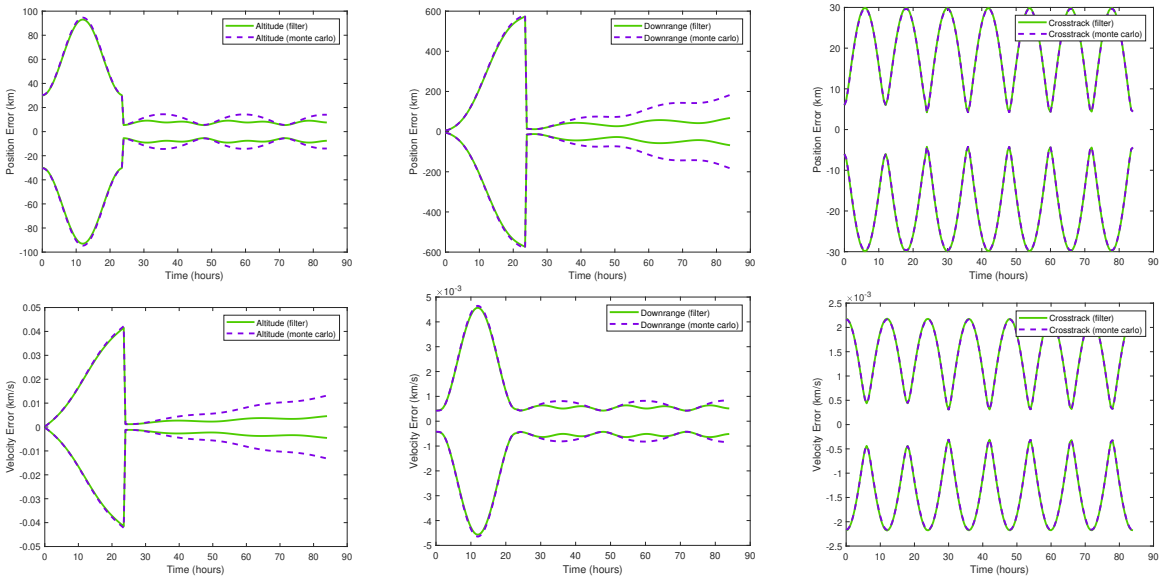


Fig. F.2: EKF Measurement at 24 Hours, OBS 2 Only

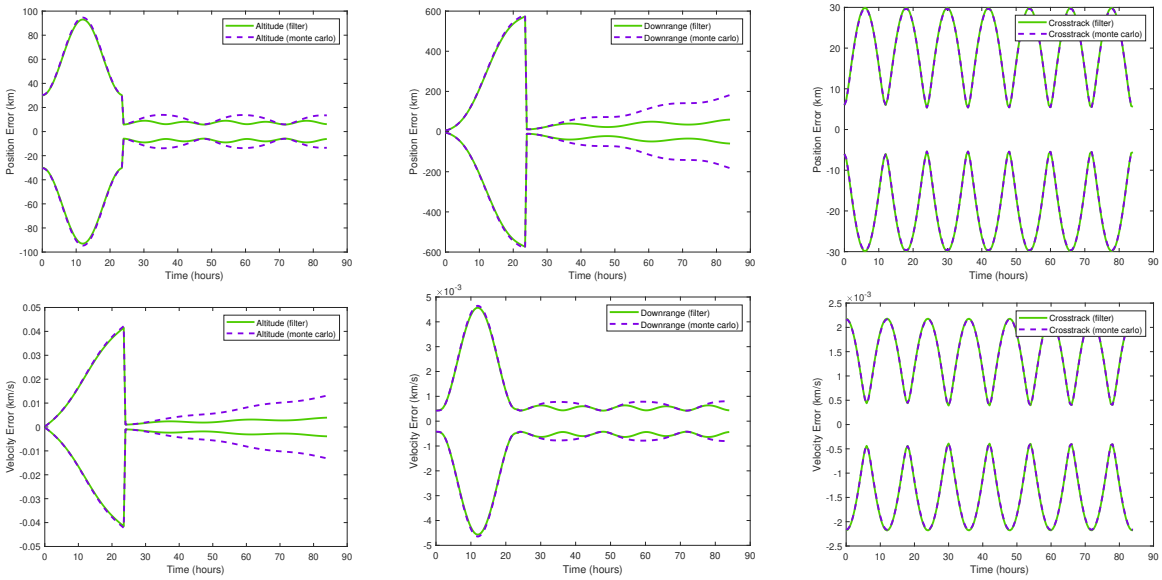


Fig. F.3: EKF Measurement at 24 Hours, OBS 3 Only

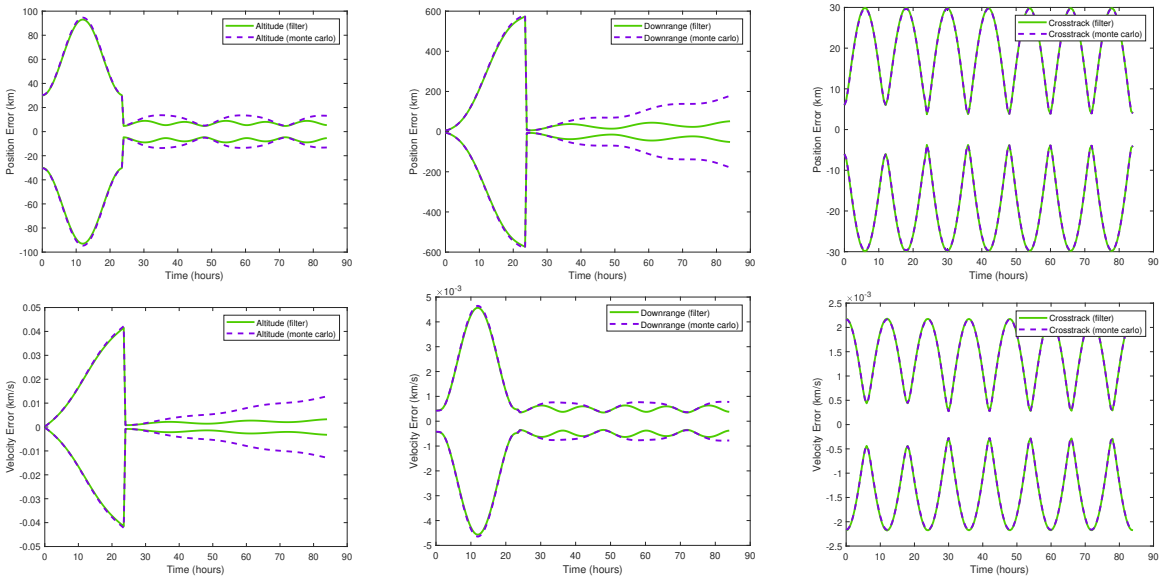


Fig. F.4: EKF Measurement at 24 Hours, All Observations

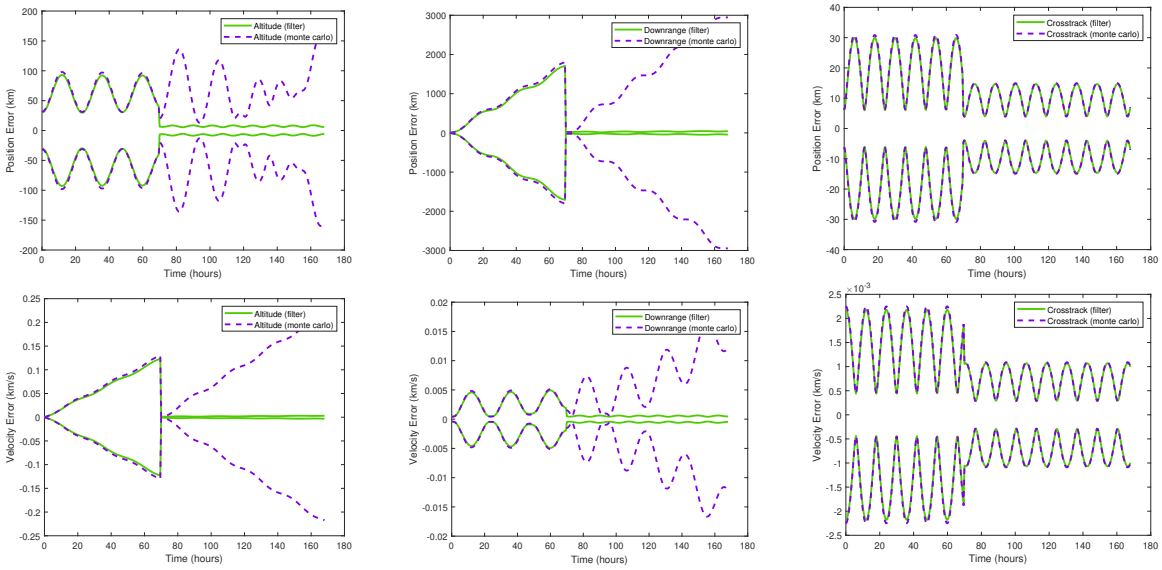


Fig. F.5: EKF Measurement at 70 Hours, OBS 1 Only

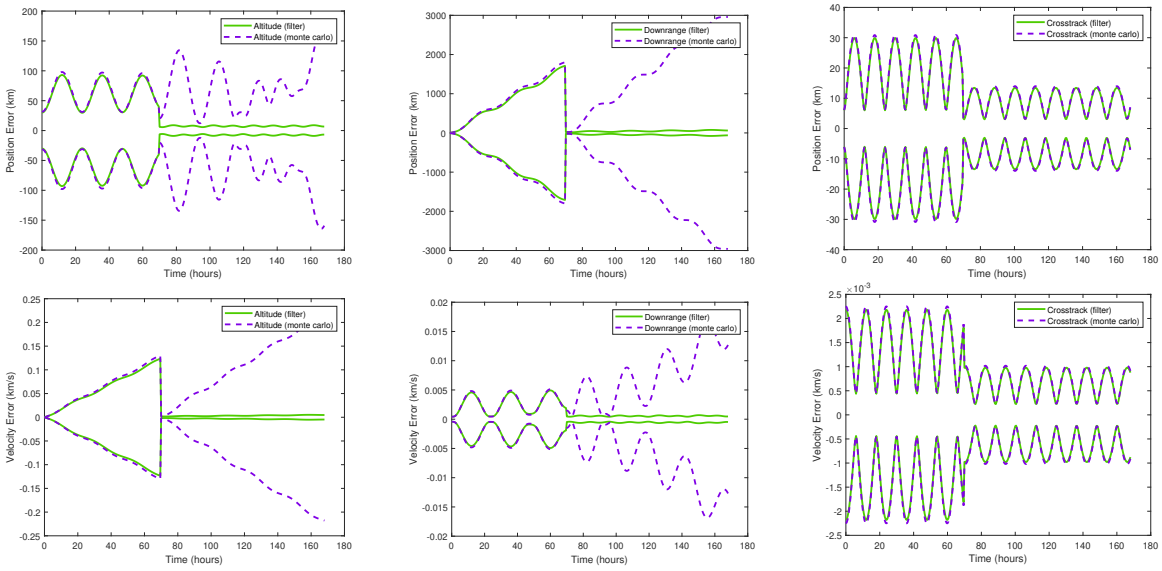


Fig. F.6: EKF Measurement at 70 Hours, OBS 2 Only

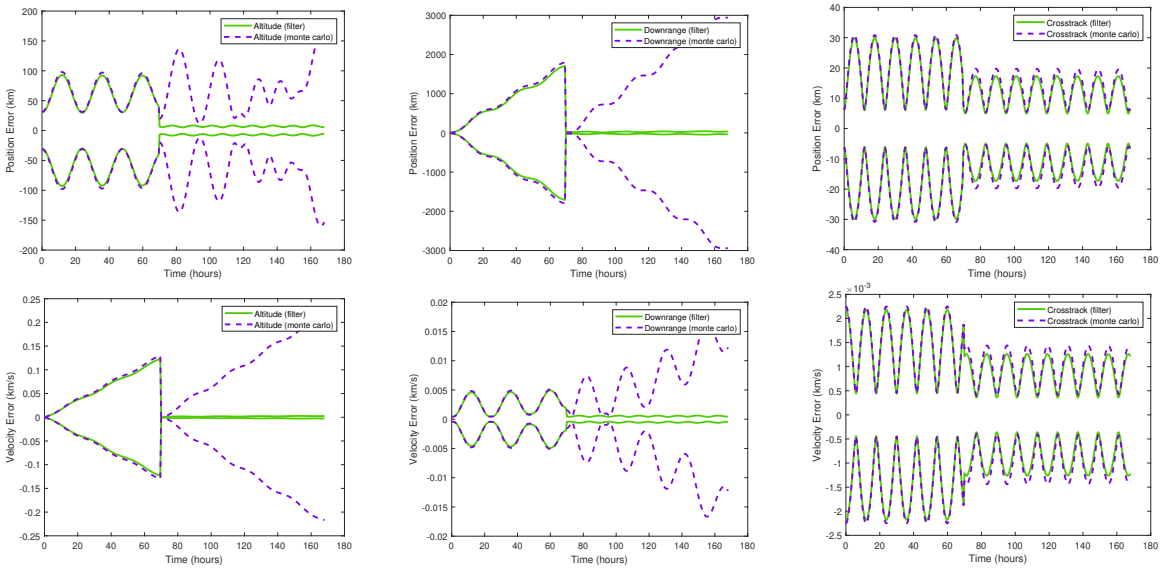


Fig. F.7: EKF Measurement at 70 Hours, OBS 3 Only

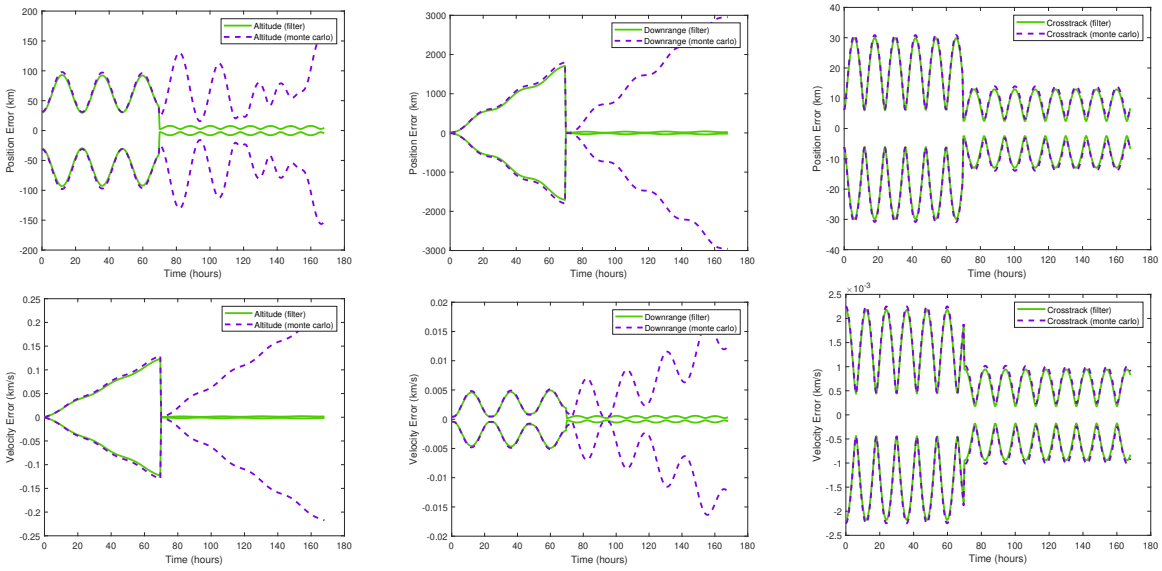


Fig. F.8: EKF Measurement at 70 Hours, All Observations

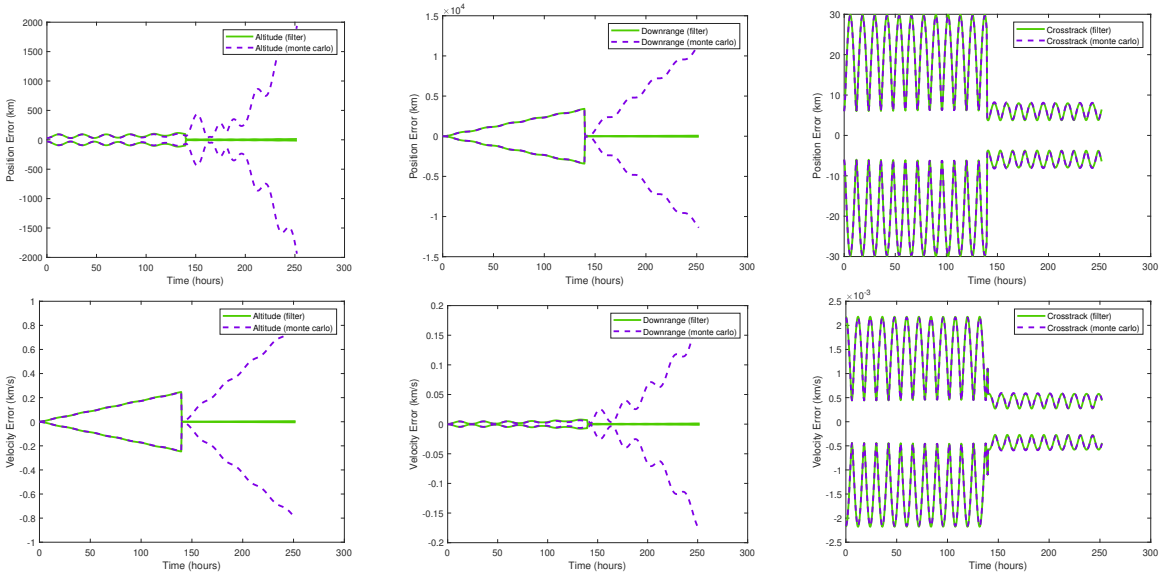


Fig. F.9: EKF Measurement at 140 Hours, OBS 1 Only

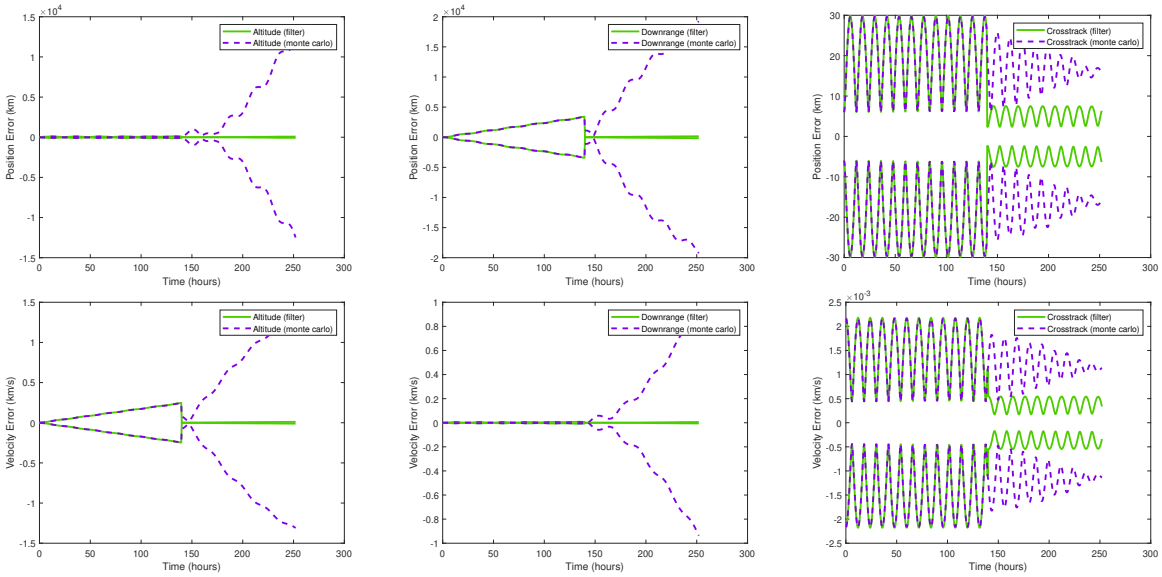


Fig. F.10: EKF Measurement at 140 Hours, OBS 2 Only

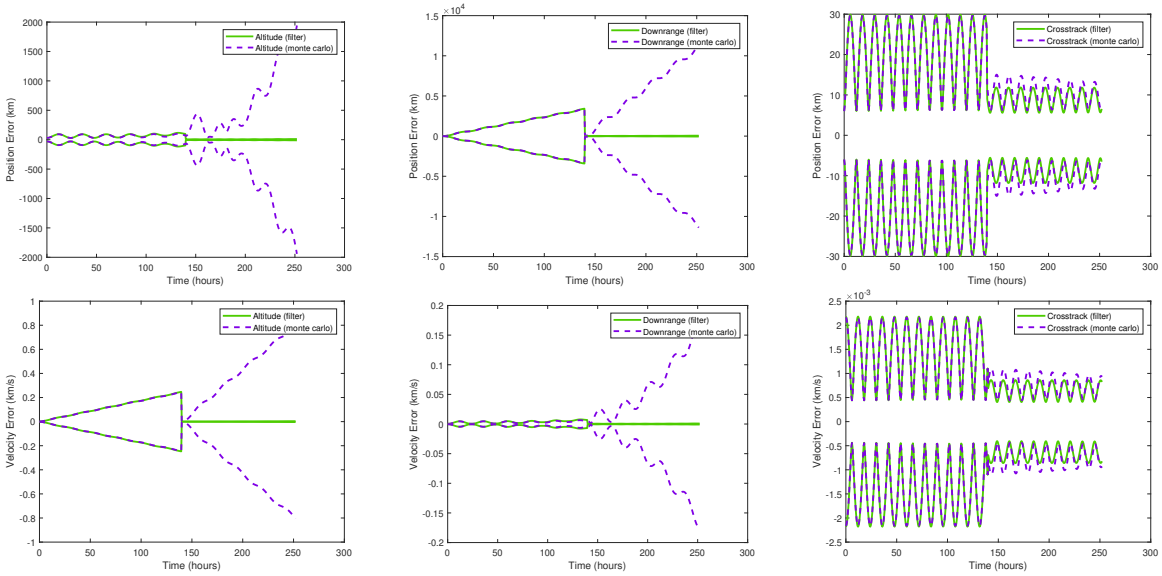


Fig. F.11: EKF Measurement at 140 Hours, OBS 3 Only

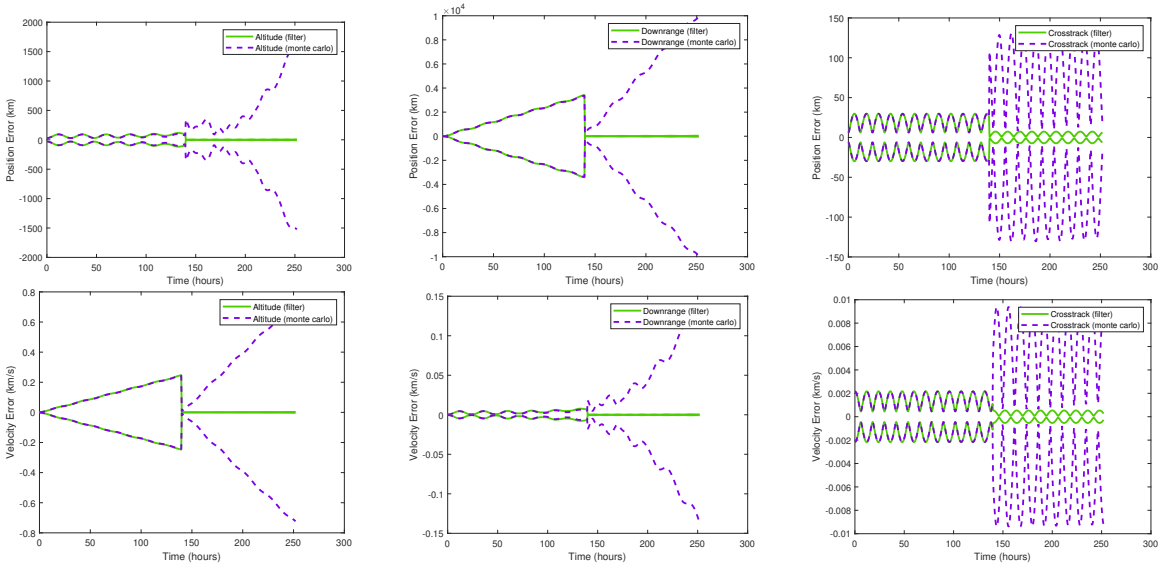


Fig. F.12: EKF Measurement at 140 Hours, All Observations

F.2 HYBRID KALMAN-PARTICLE FILTER

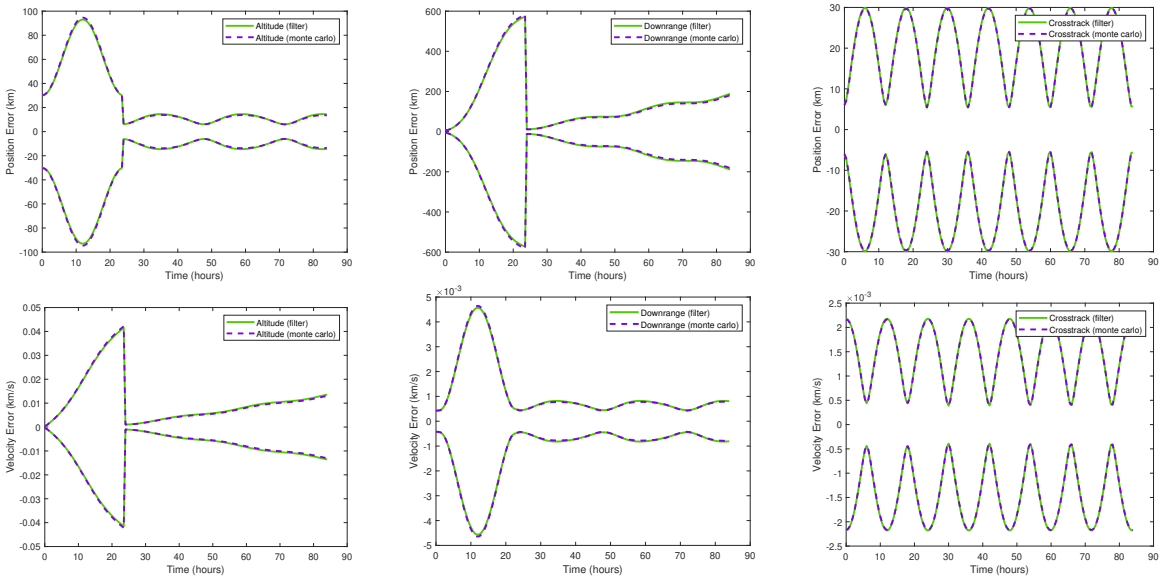


Fig. F.13: HKF Measurement at 24 Hours, OBS 1 Only

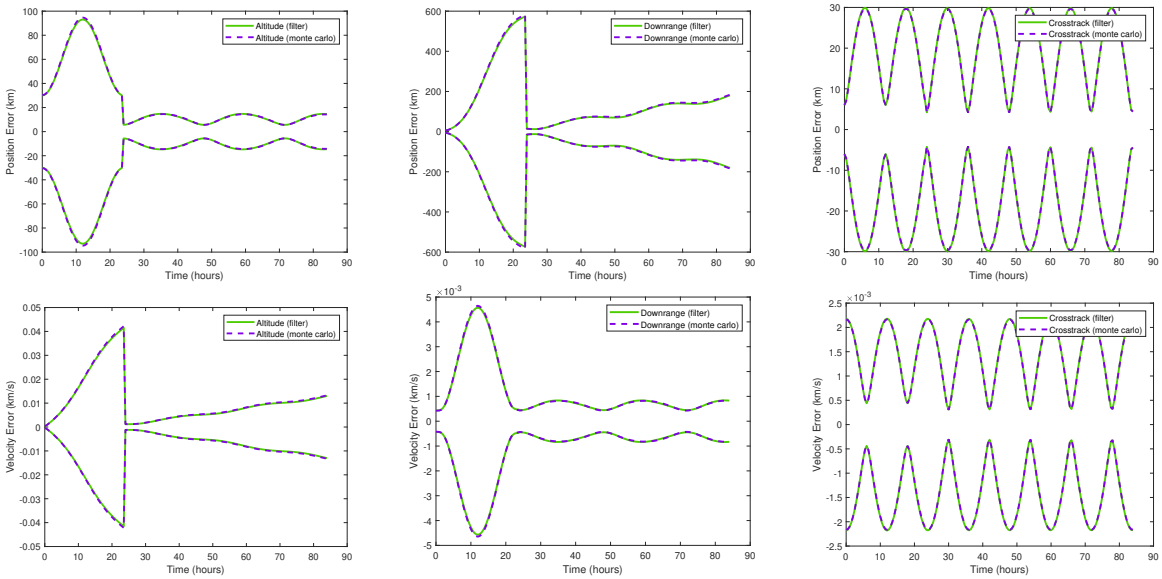


Fig. F.14: HKF Measurement at 24 Hours, OBS 2 Only

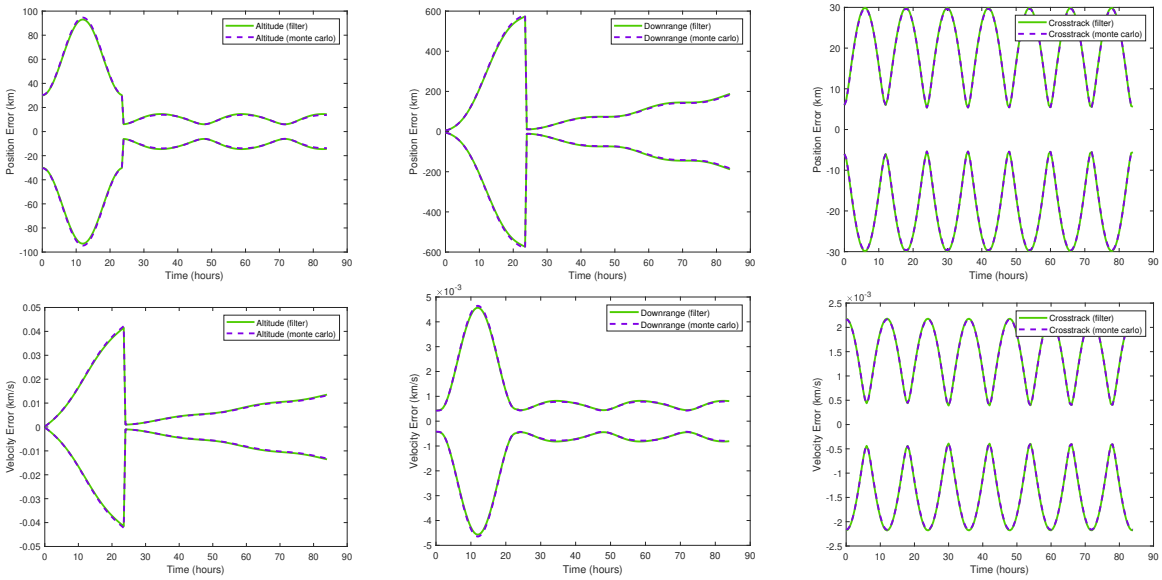


Fig. F.15: HKF Measurement at 24 Hours, OBS 3 Only

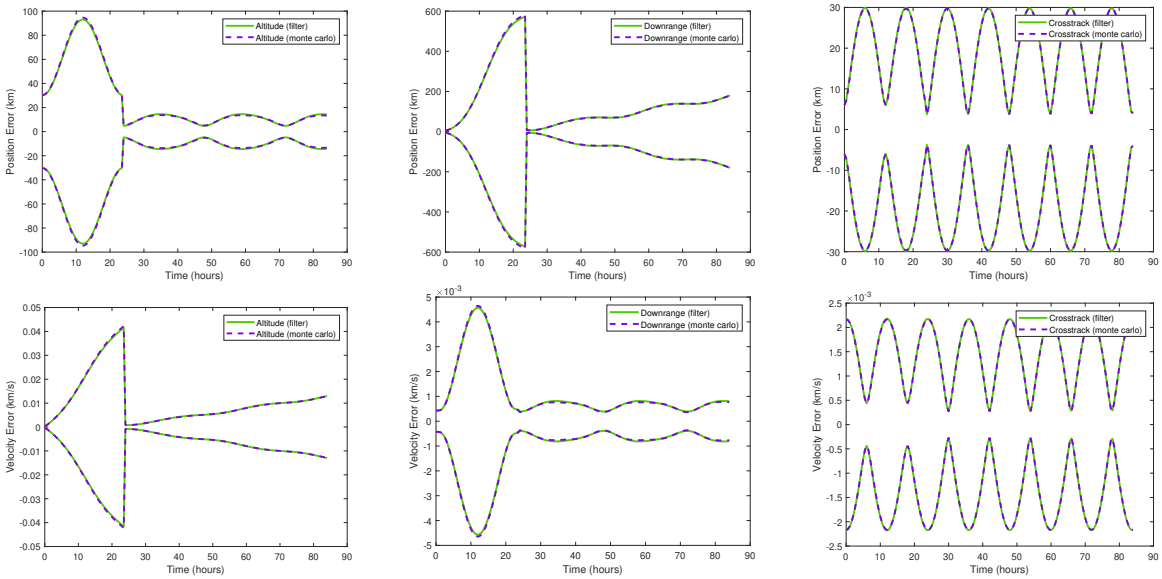


Fig. F.16: HKF Measurement at 24 Hours, All Observations

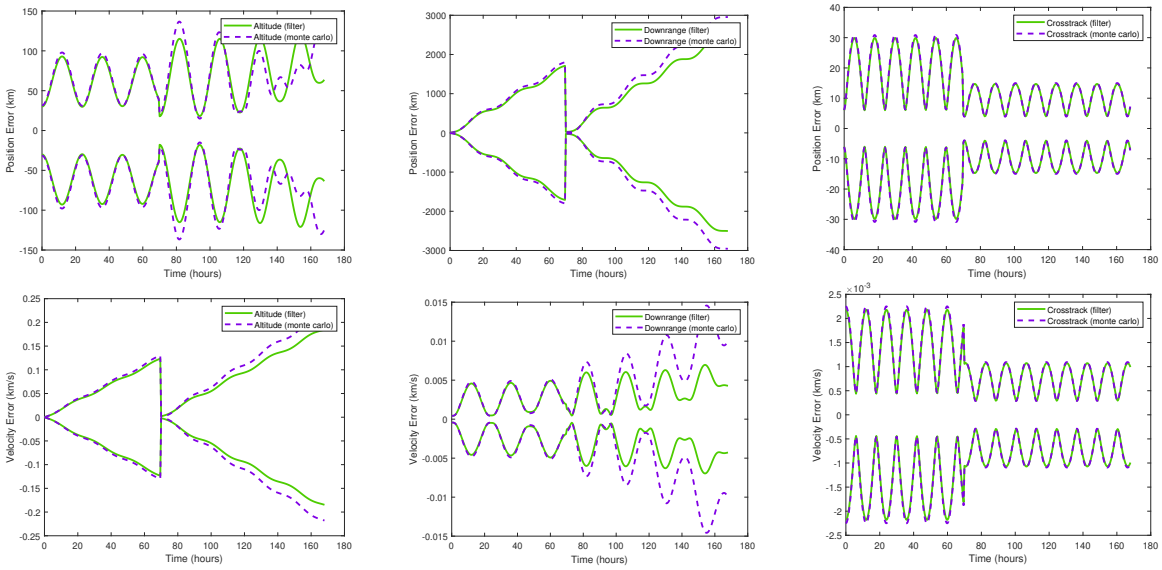


Fig. F.17: HKF Measurement at 70 Hours, OBS 1 Only

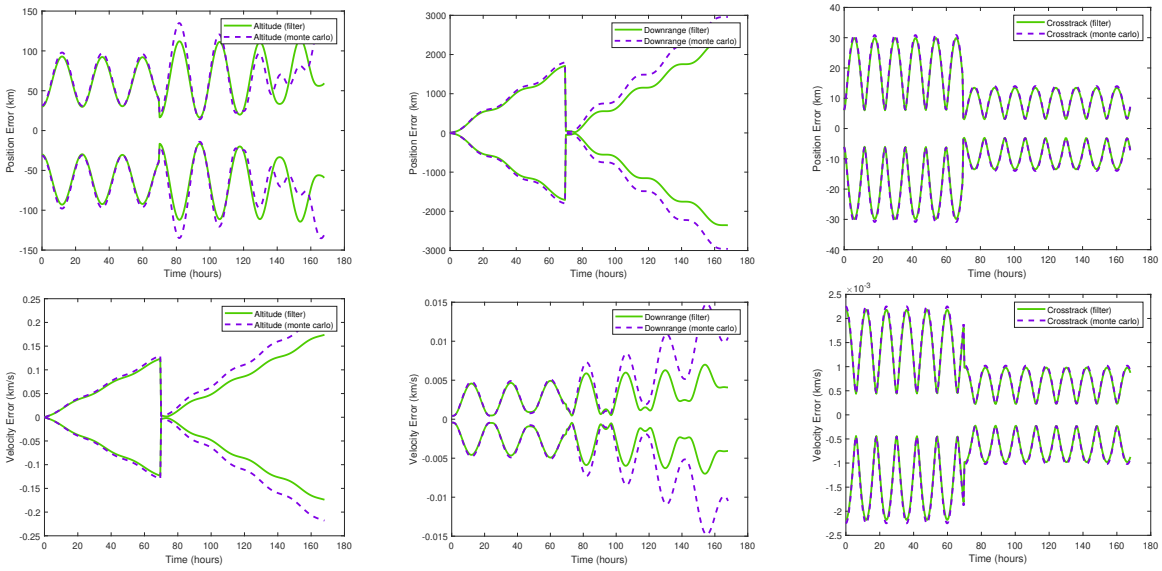


Fig. F.18: HKF Measurement at 70 Hours, OBS 2 Only

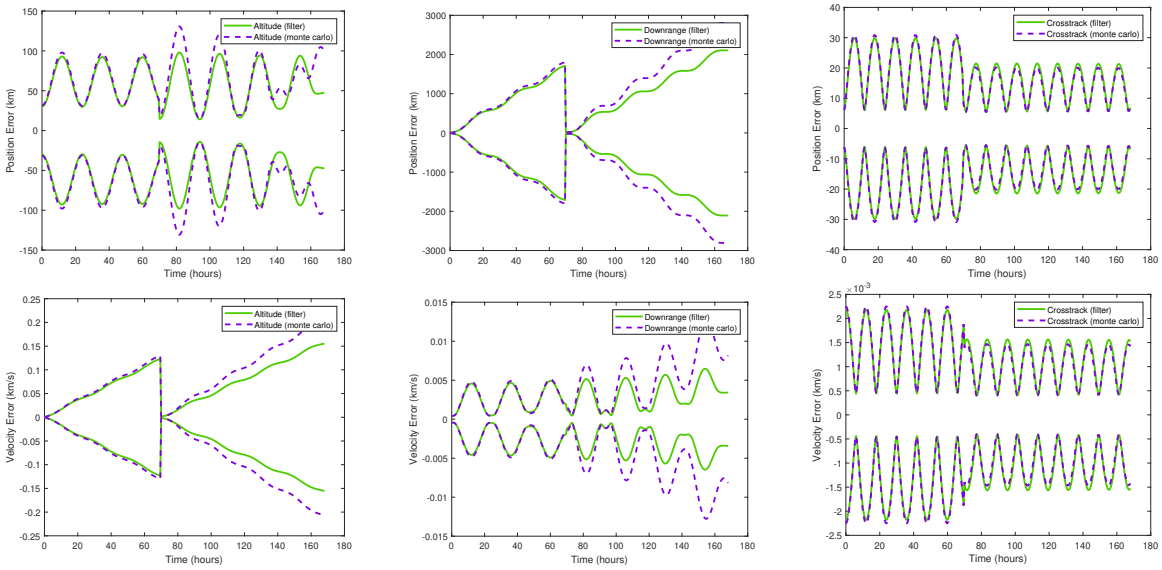


Fig. F.19: HKF Measurement at 70 Hours, OBS 3 Only

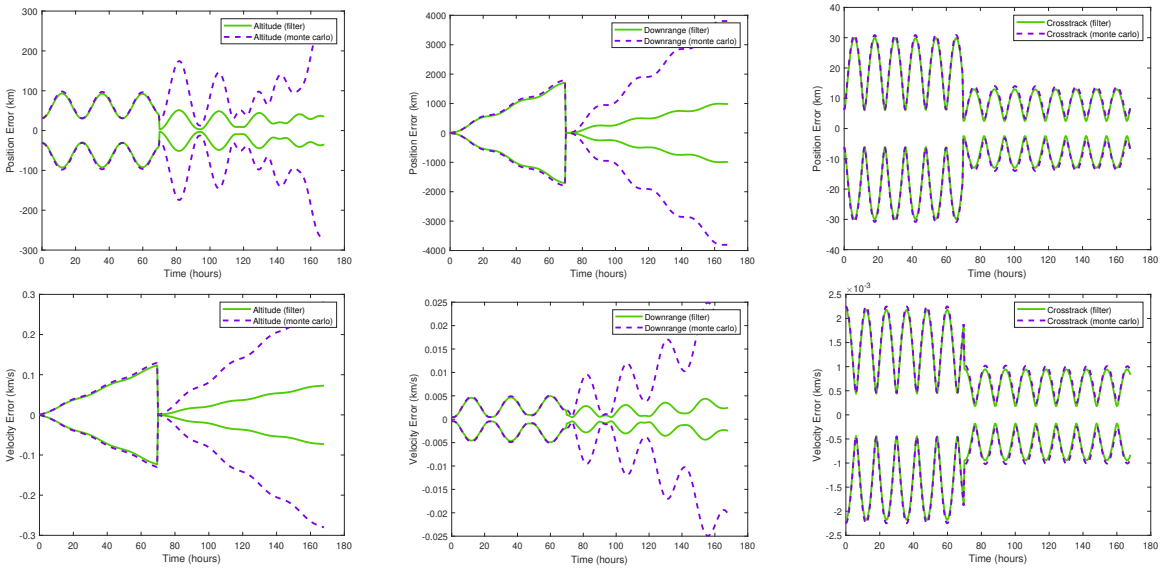


Fig. F.20: HKF Measurement at 70 Hours, All Observations

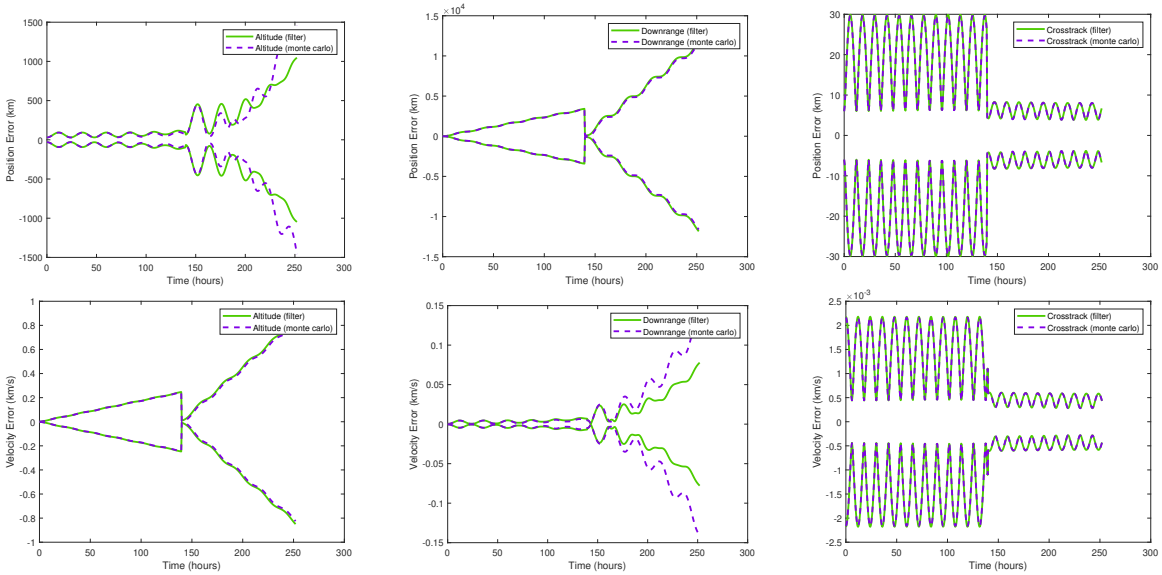


Fig. F.21: HKF Measurement at 140 Hours, OBS 1 Only

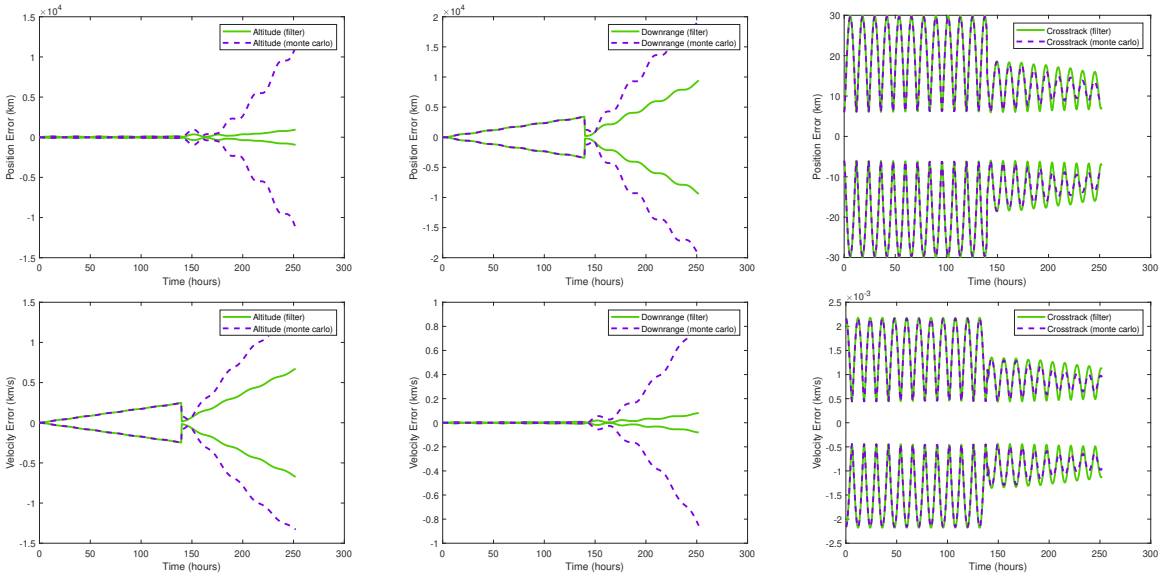


Fig. F.22: HKF Measurement at 140 Hours, OBS 2 Only

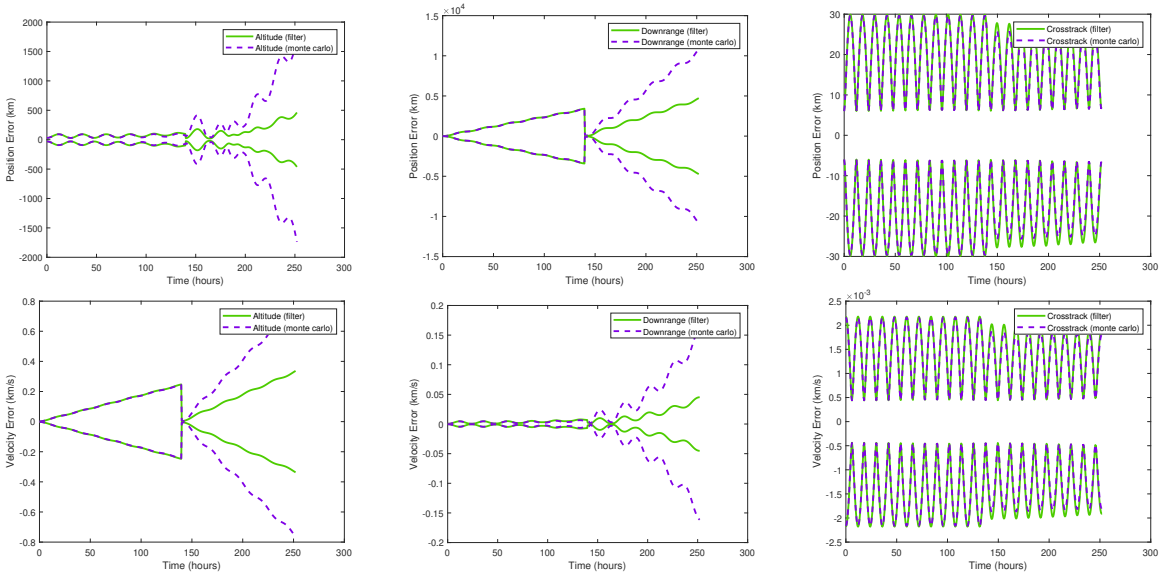


Fig. F.23: HKF Measurement at 140 Hours, OBS 3 Only

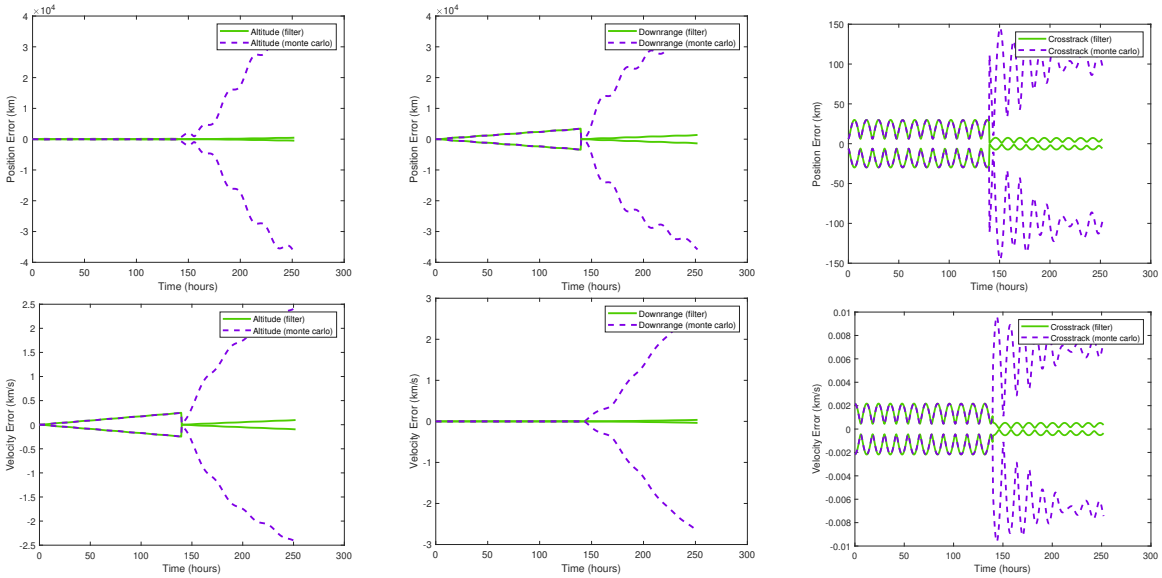


Fig. F.24: HKF Measurement at 140 Hours, All Observations

F.3 UNSCENTED KALMAN FILTER

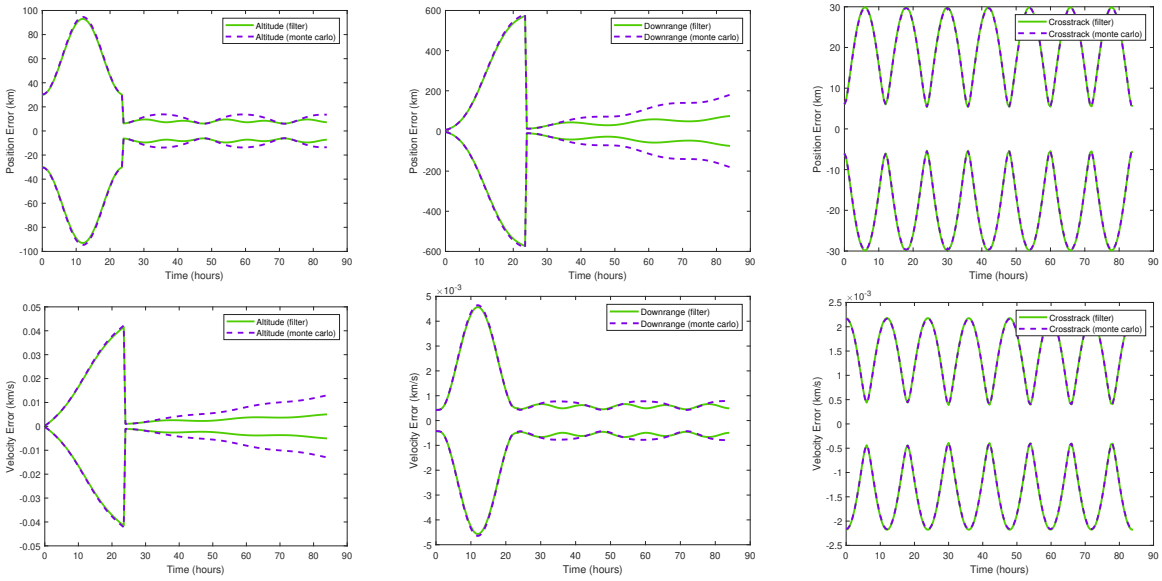


Fig. F.25: UKF Measurement at 24 Hours, OBS 1 Only

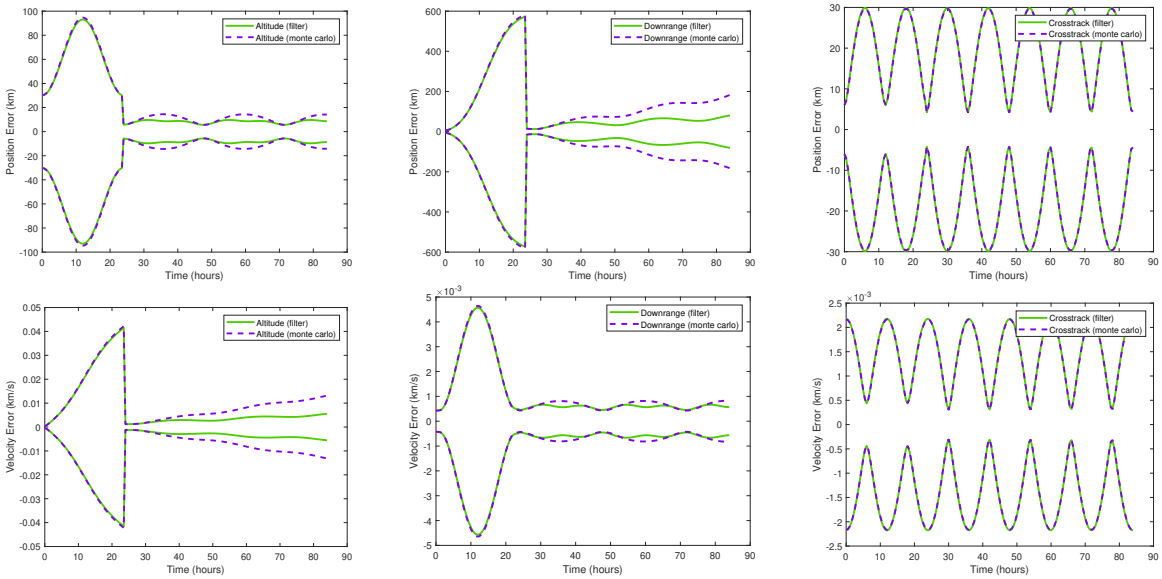


Fig. F.26: UKF Measurement at 24 Hours, OBS 2 Only

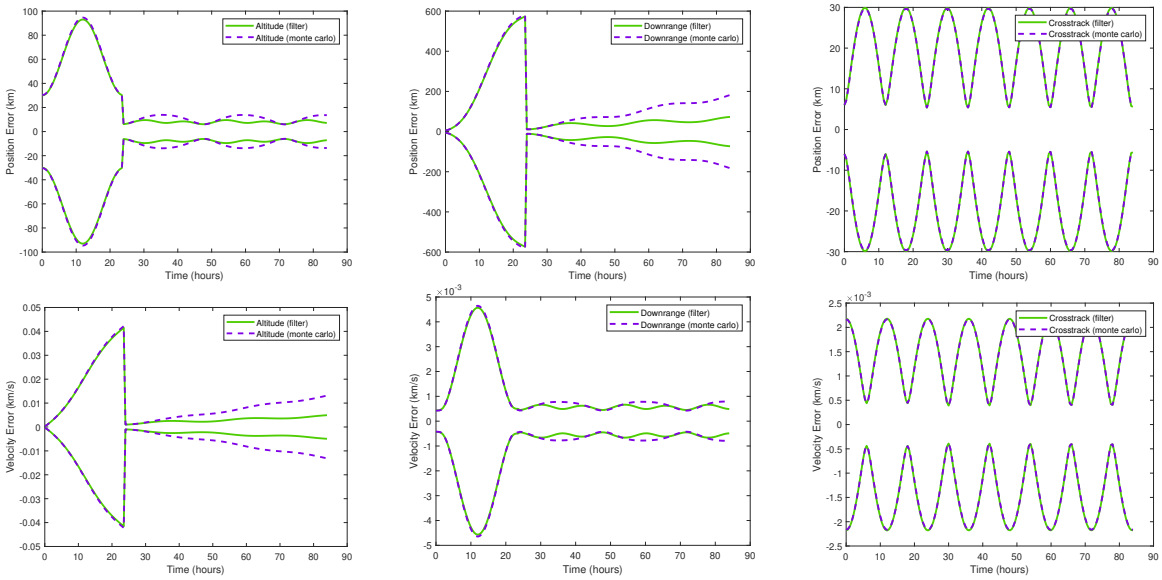


Fig. F.27: UKF Measurement at 24 Hours, OBS 3 Only

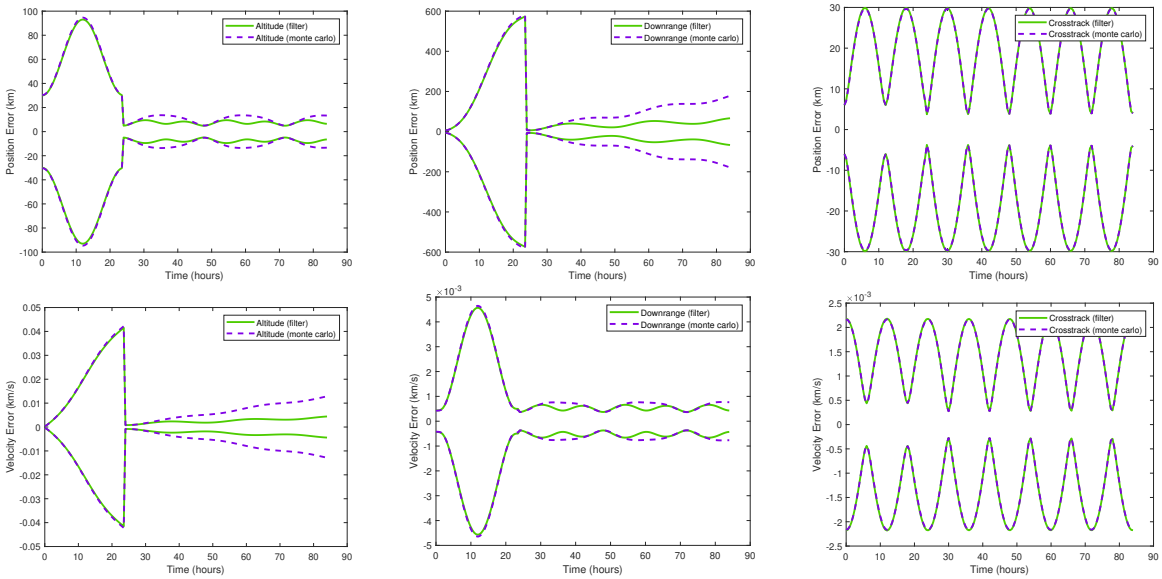


Fig. F.28: UKF Measurement at 24 Hours, All Observations

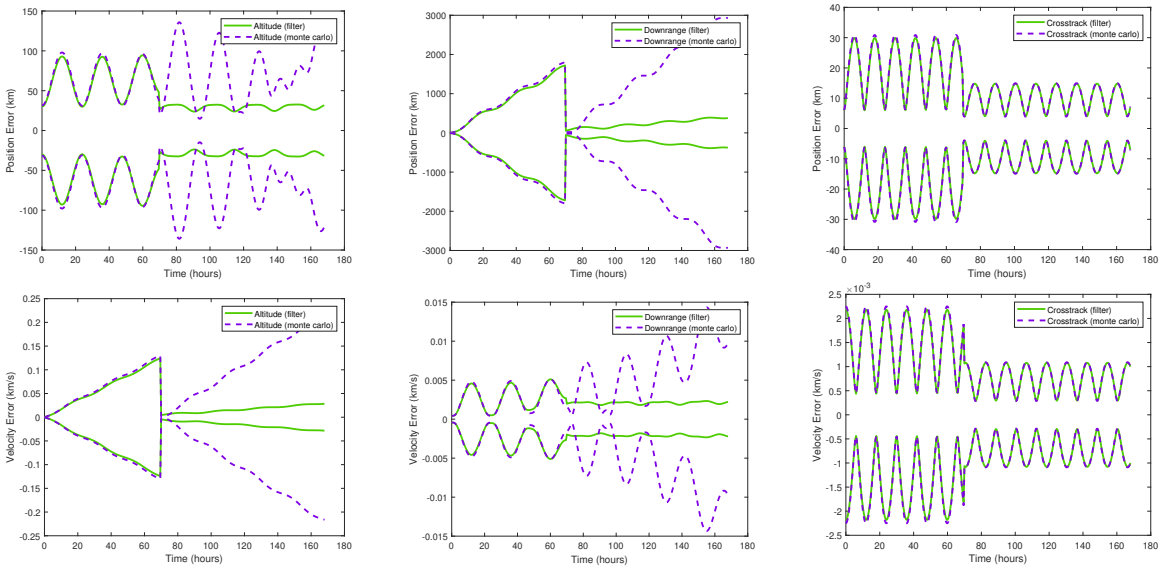


Fig. F.29: UKF Measurement at 70 Hours, OBS 1 Only

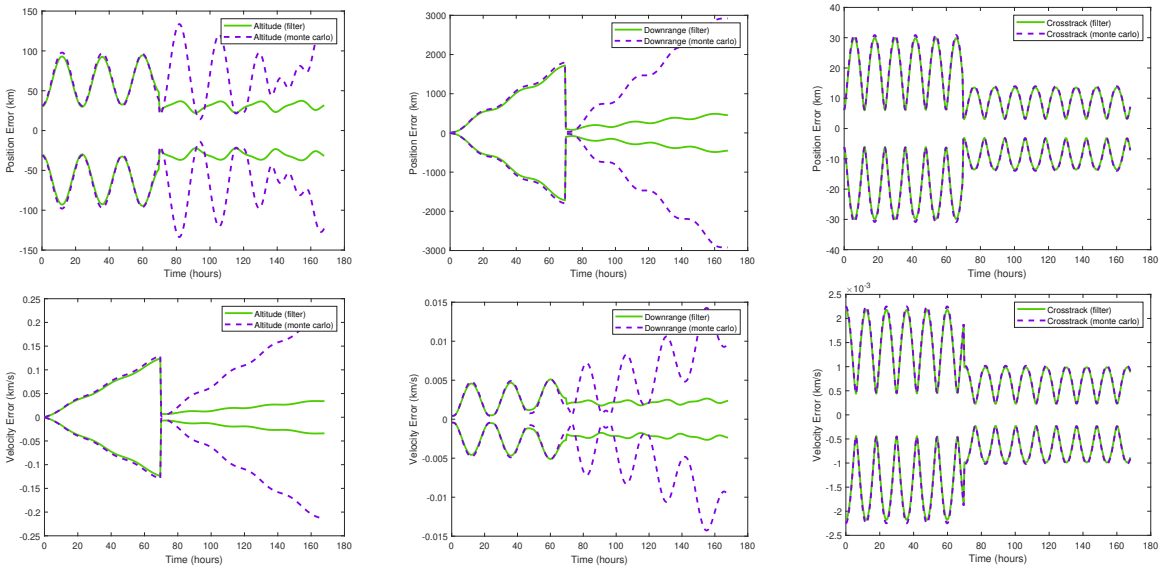


Fig. F.30: UKF Measurement at 70 Hours, OBS 2 Only

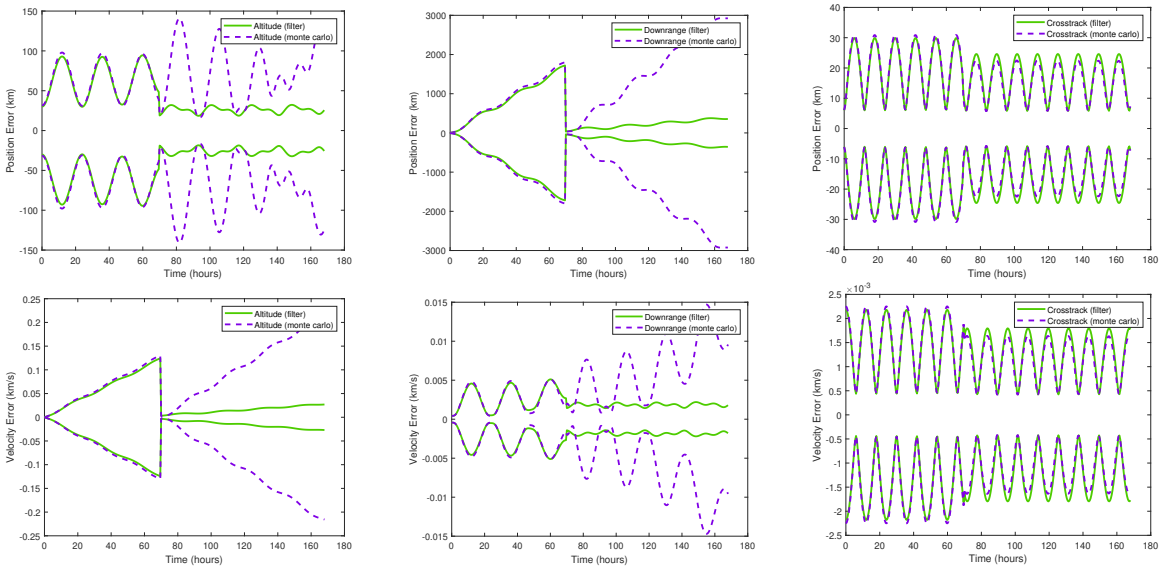


Fig. F.31: UKF Measurement at 70 Hours, OBS 3 Only

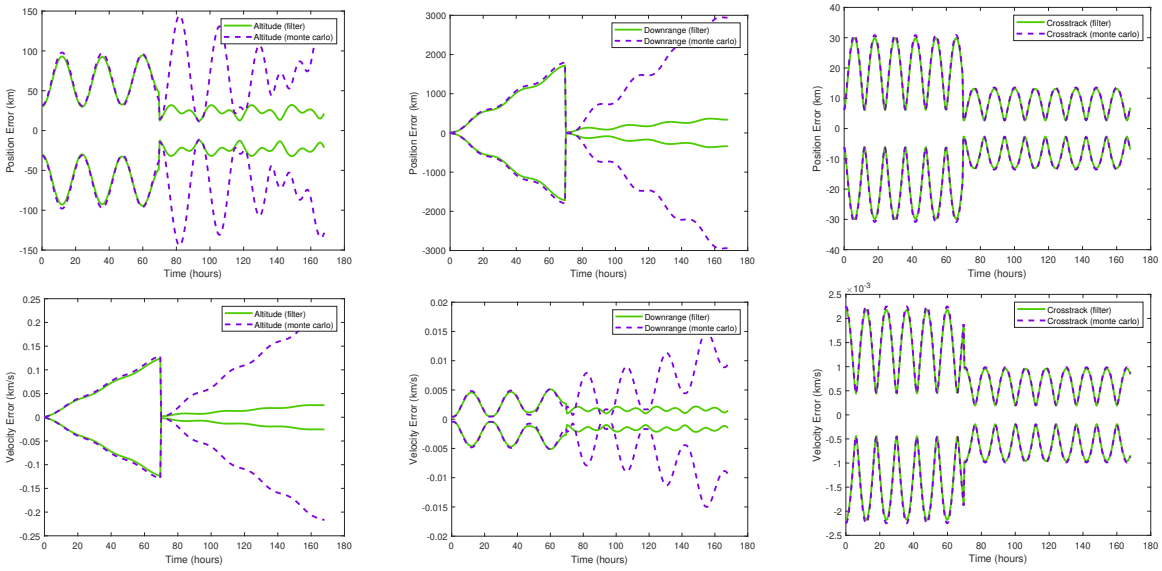


Fig. F.32: UKF Measurement at 70 Hours, All Observations

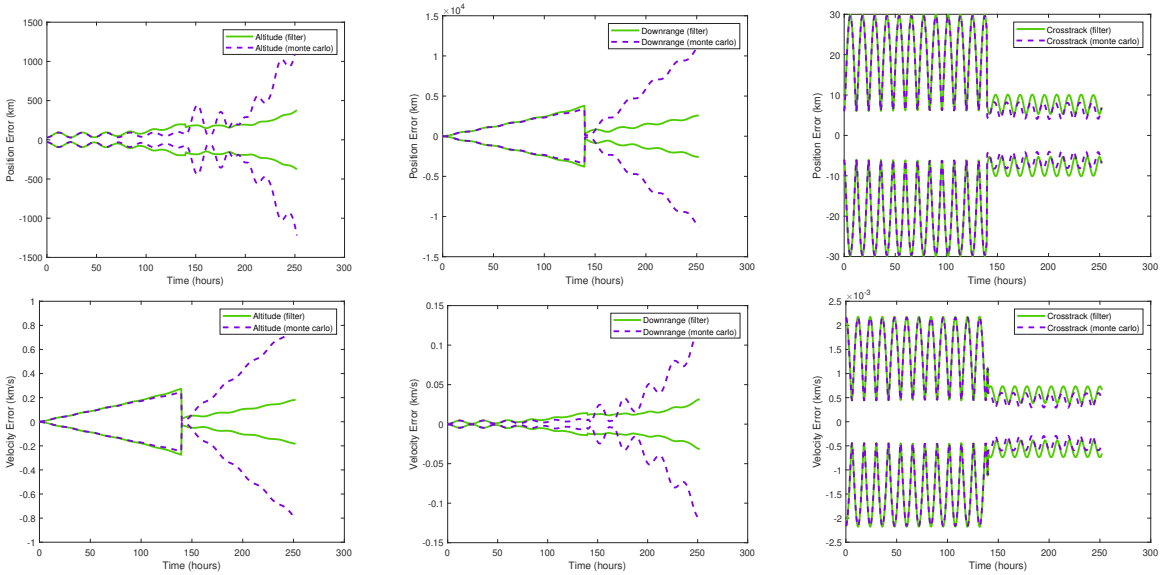


Fig. F.33: UKF Measurement at 140 Hours, OBS 1 Only

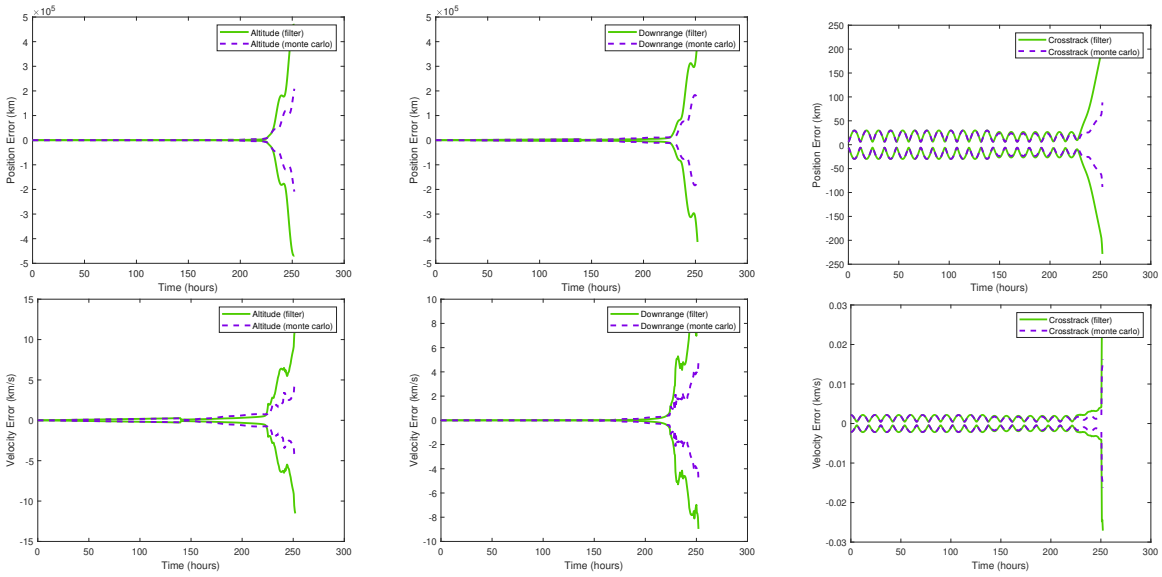


Fig. F.34: UKF Measurement at 140 Hours, OBS 2 Only

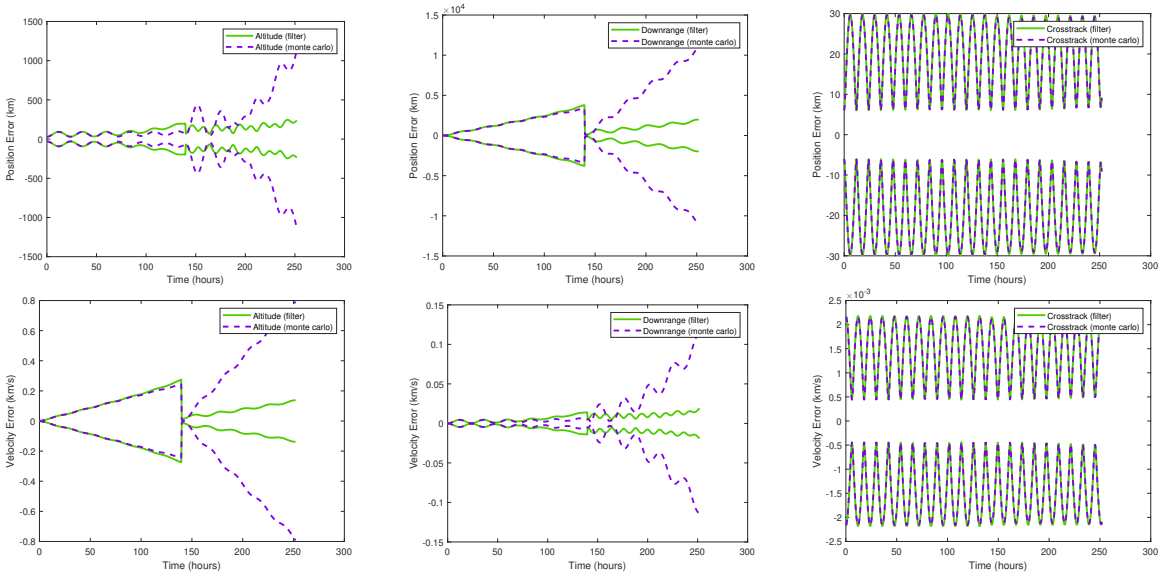


Fig. F.35: UKF Measurement at 140 Hours, OBS 3 Only

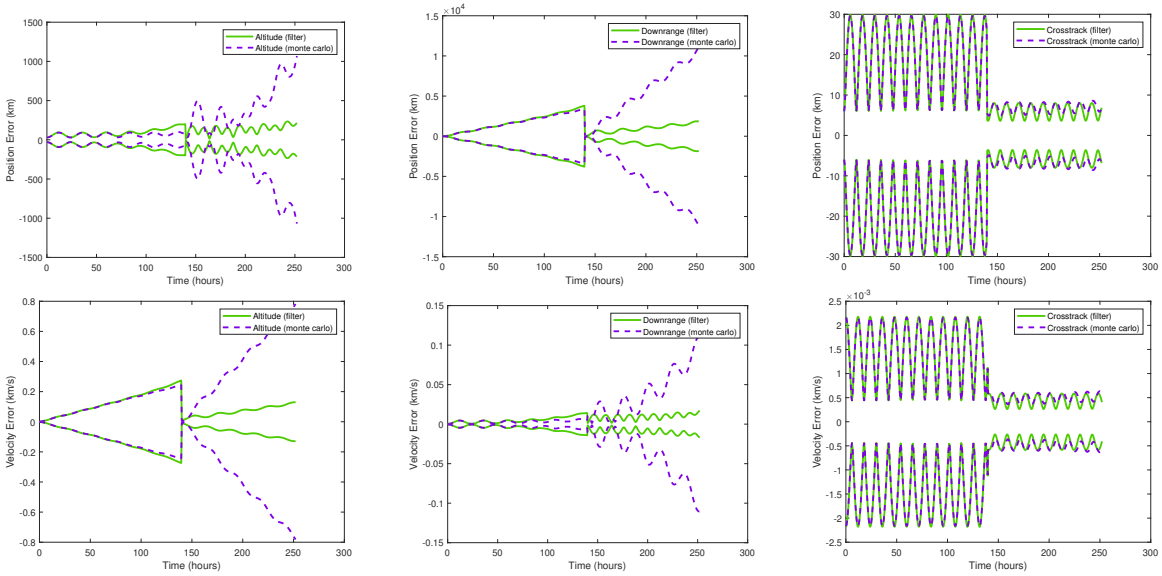


Fig. F.36: UKF Measurement at 140 Hours, All Observations

F.4 GAUSSIAN MIXTURE MODEL FILTER

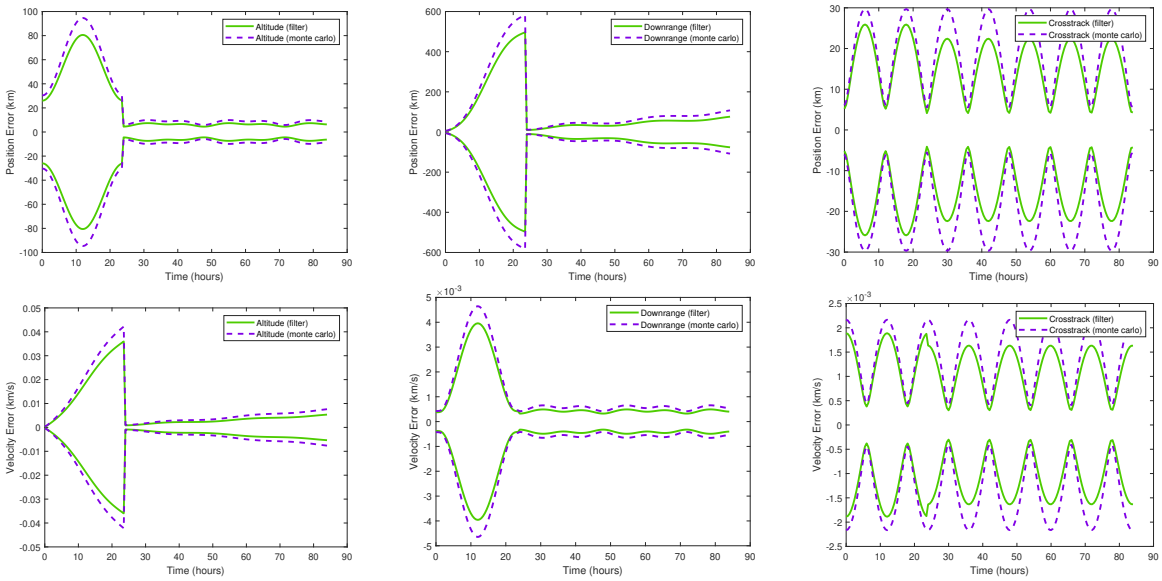


Fig. F.37: GMM Measurement at 24 Hours, OBS 1 Only

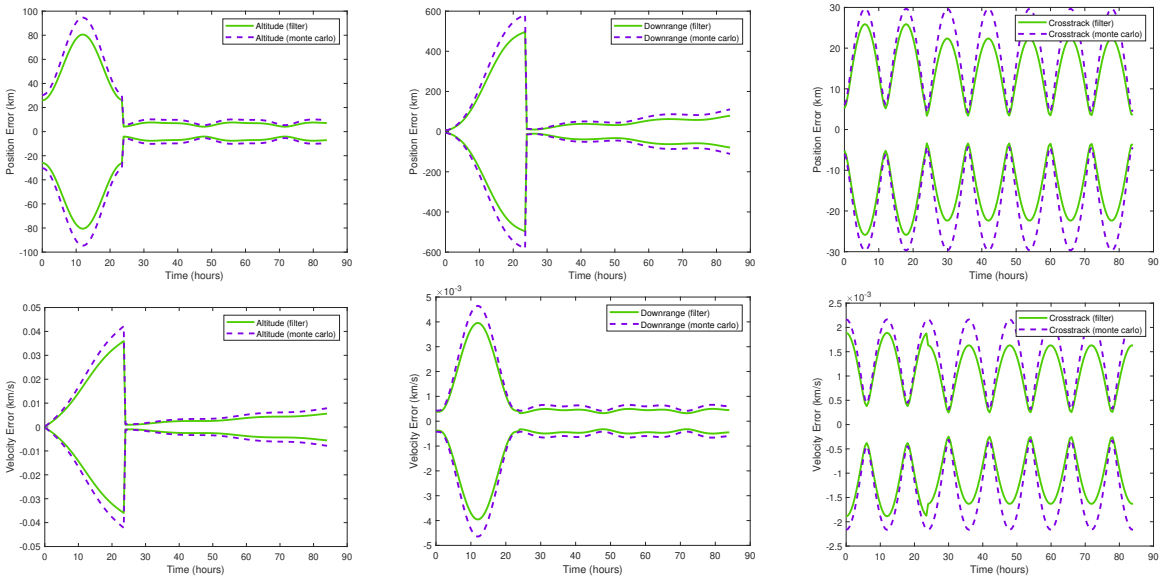


Fig. F.38: GMM Measurement at 24 Hours, OBS 2 Only

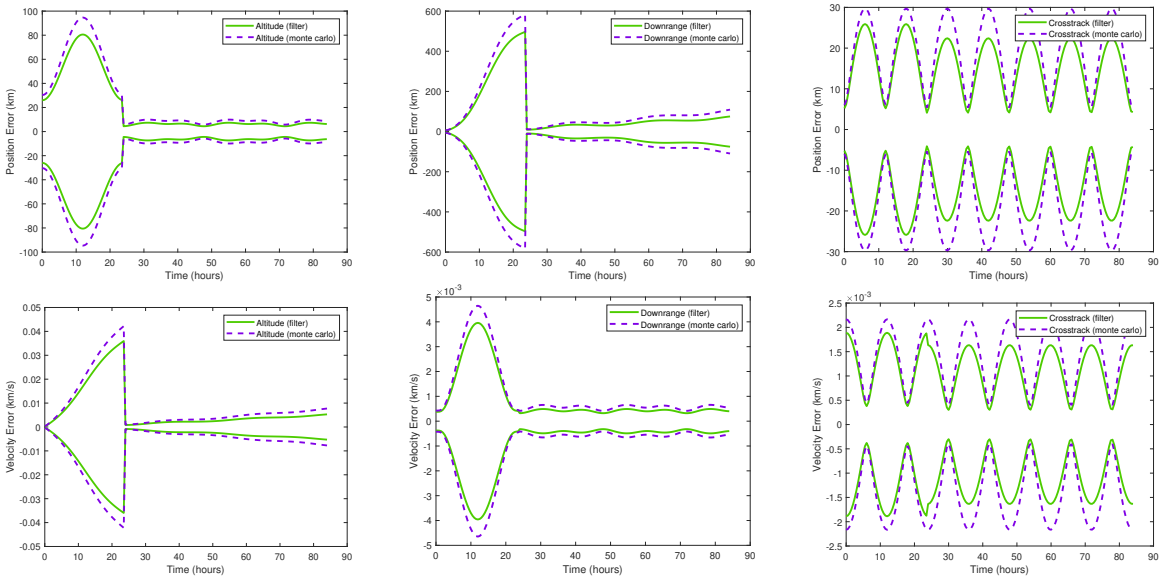


Fig. F.39: GMM Measurement at 24 Hours, OBS 3 Only

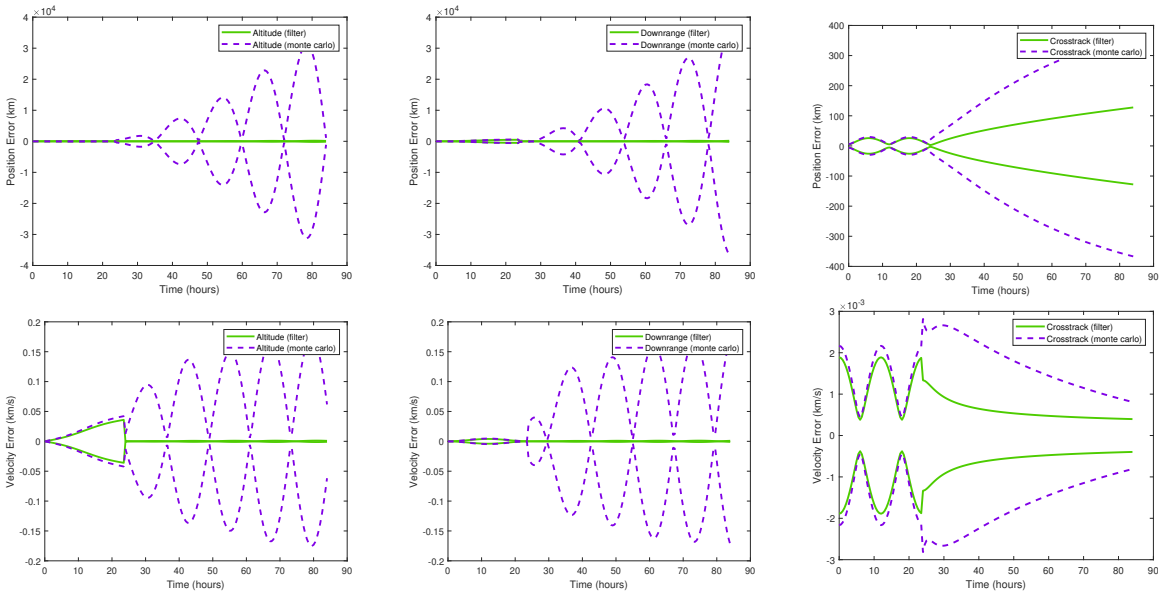


Fig. F.40: GMM Measurement at 24 Hours, All Observations

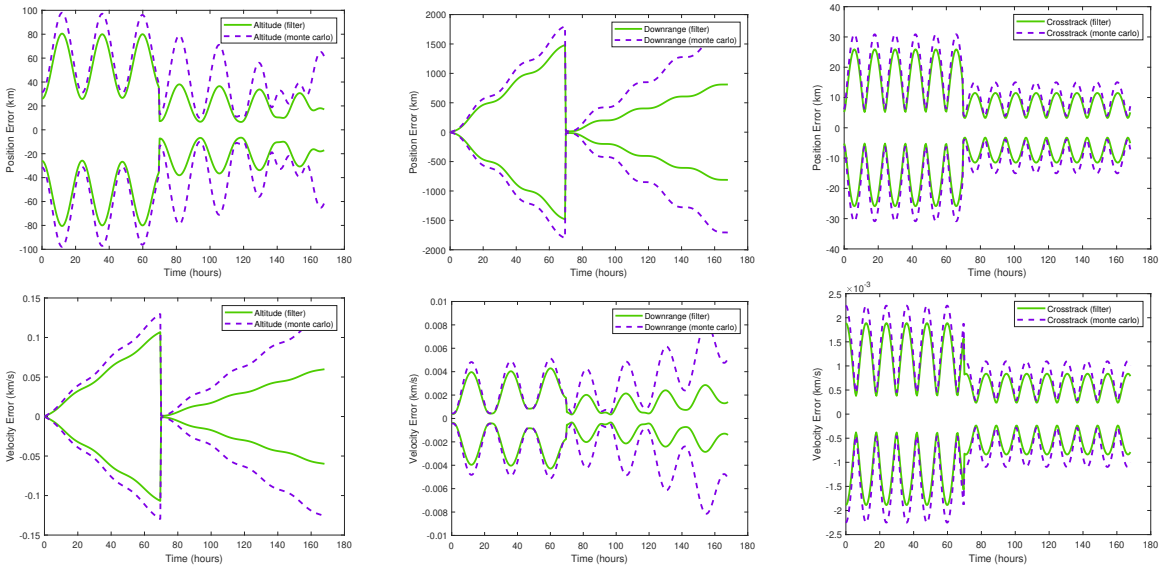


Fig. F.41: GMM Measurement at 70 Hours, OBS 1 Only

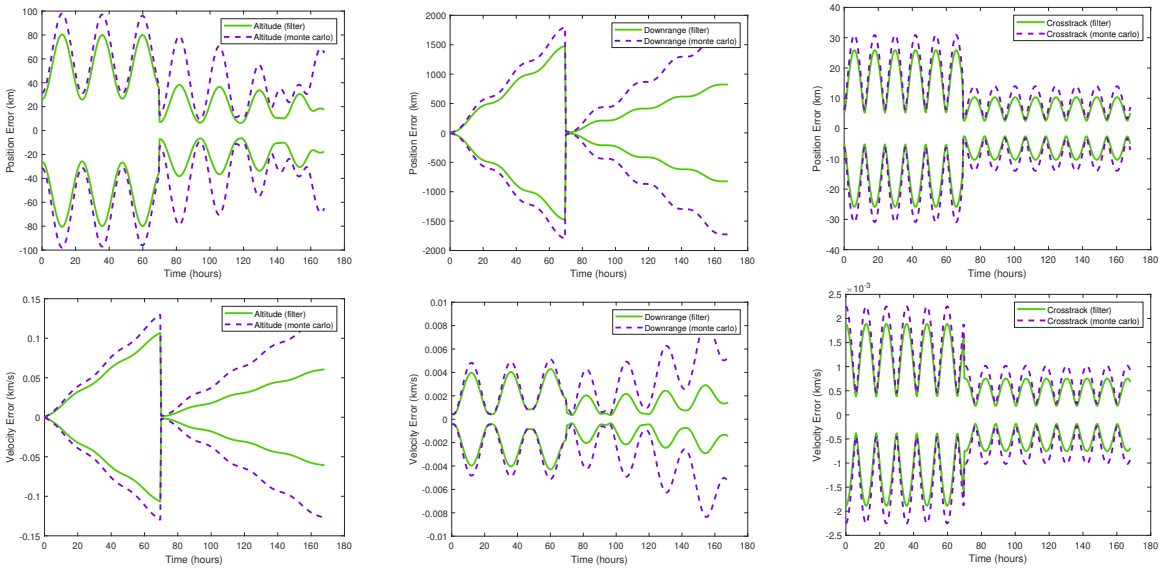


Fig. F.42: GMM Measurement at 70 Hours, OBS 2 Only

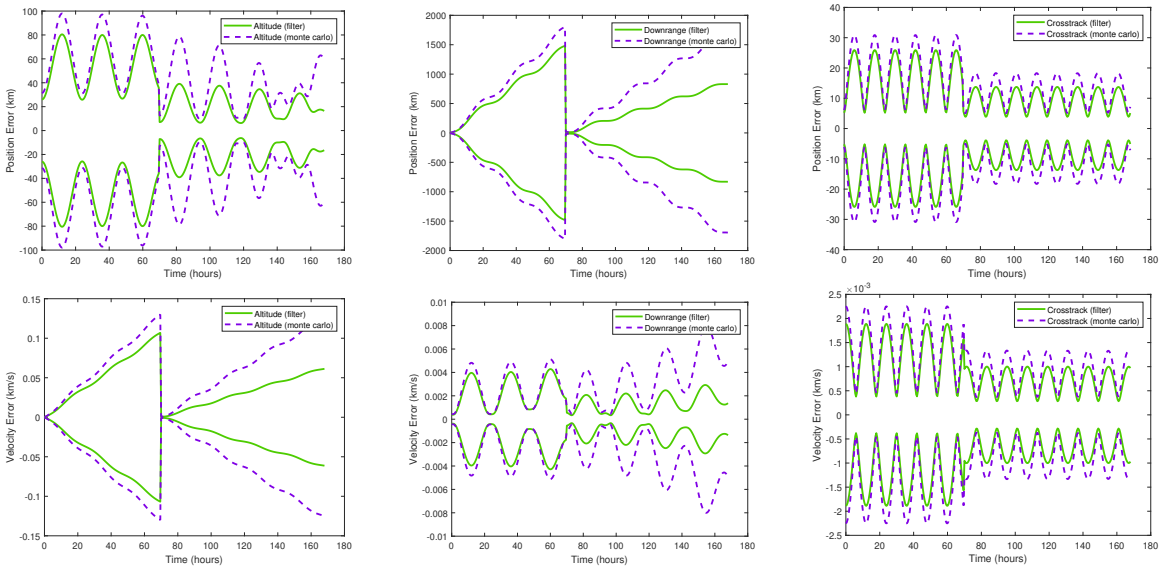


Fig. F.43: GMM Measurement at 70 Hours, OBS 3 Only

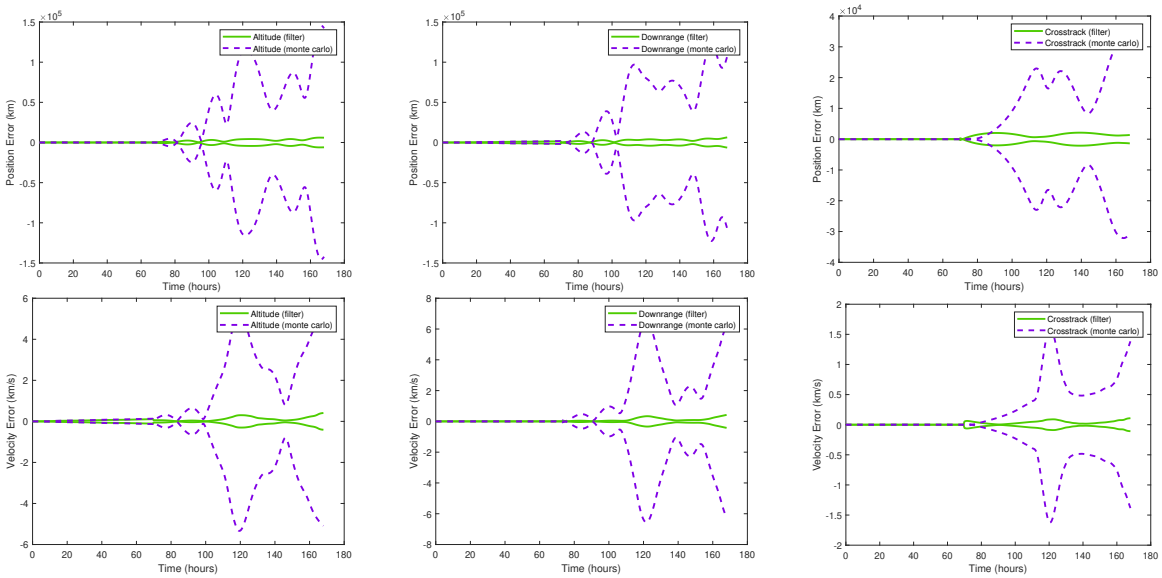


Fig. F.44: GMM Measurement at 70 Hours, All Observations

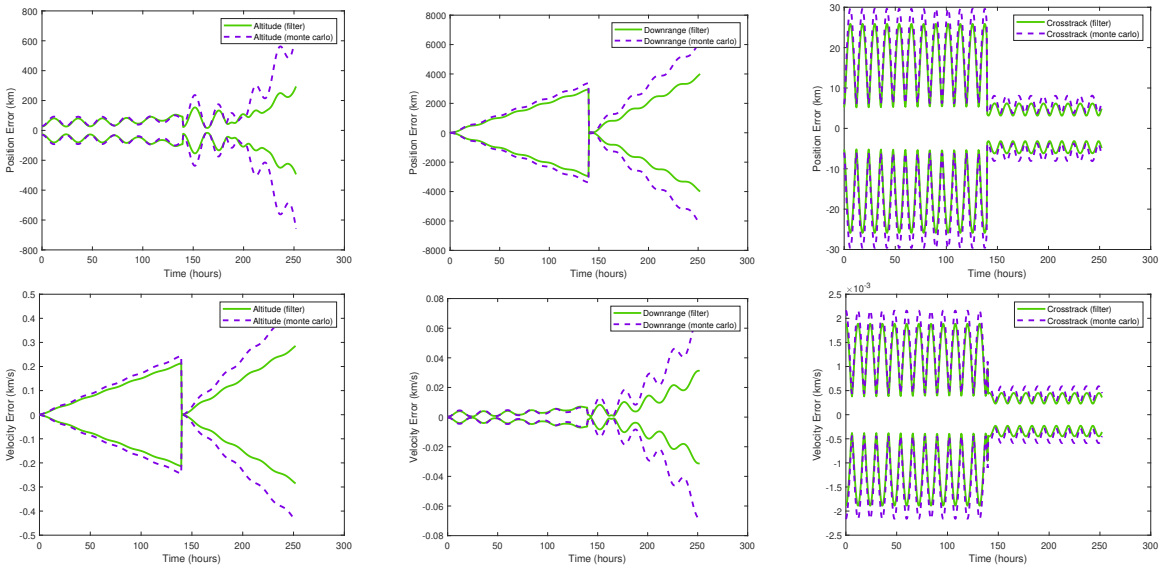


Fig. F.45: GMM Measurement at 140 Hours, OBS 1 Only

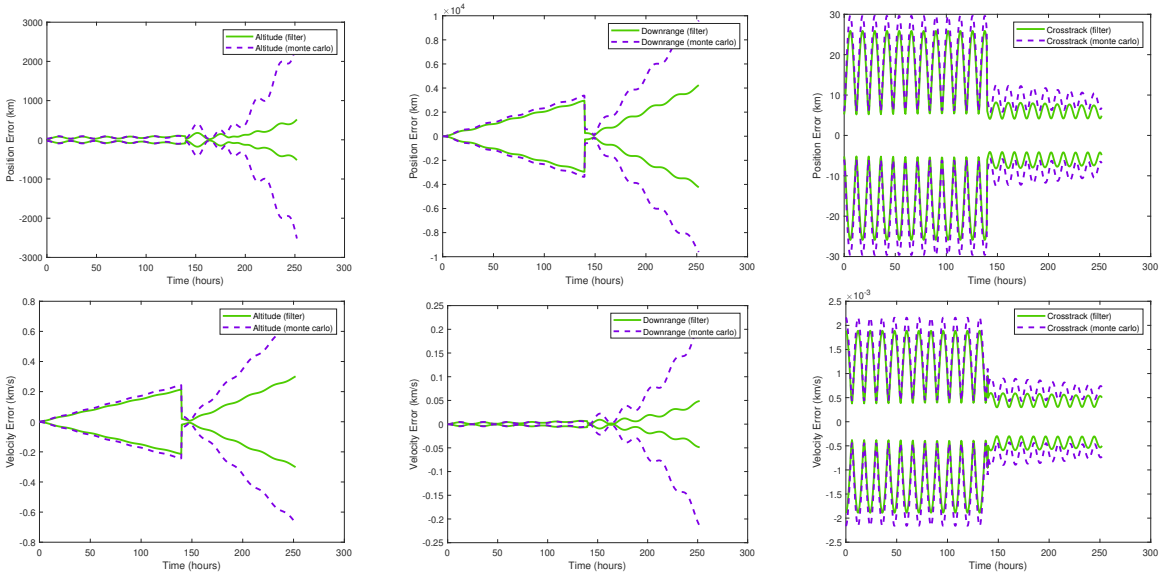


Fig. F.46: GMM Measurement at 140 Hours, OBS 2 Only

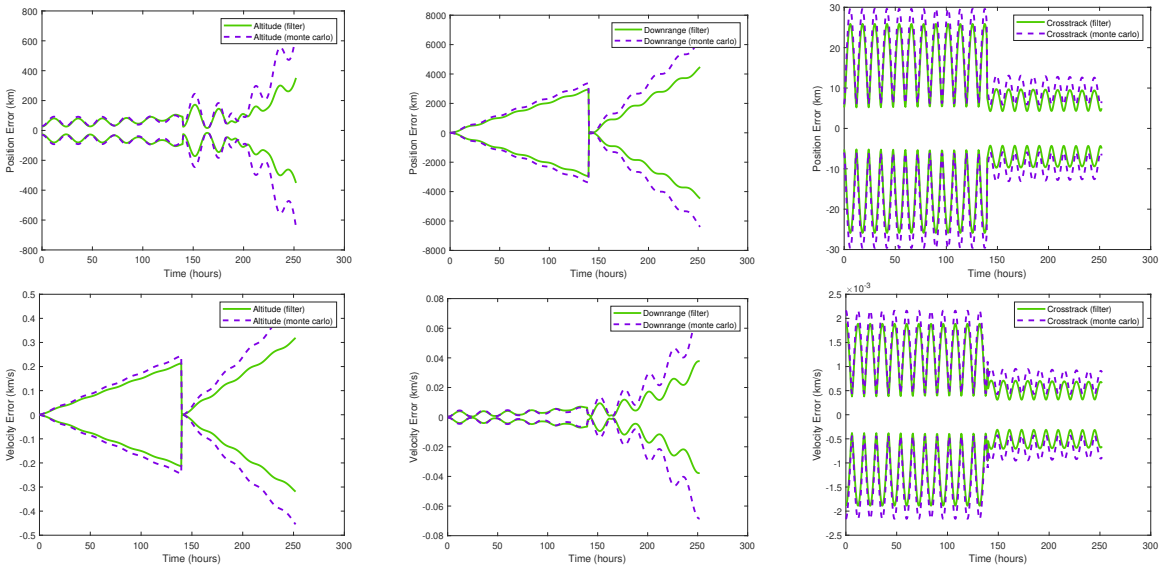


Fig. F.47: GMM Measurement at 140 Hours, OBS 3 Only

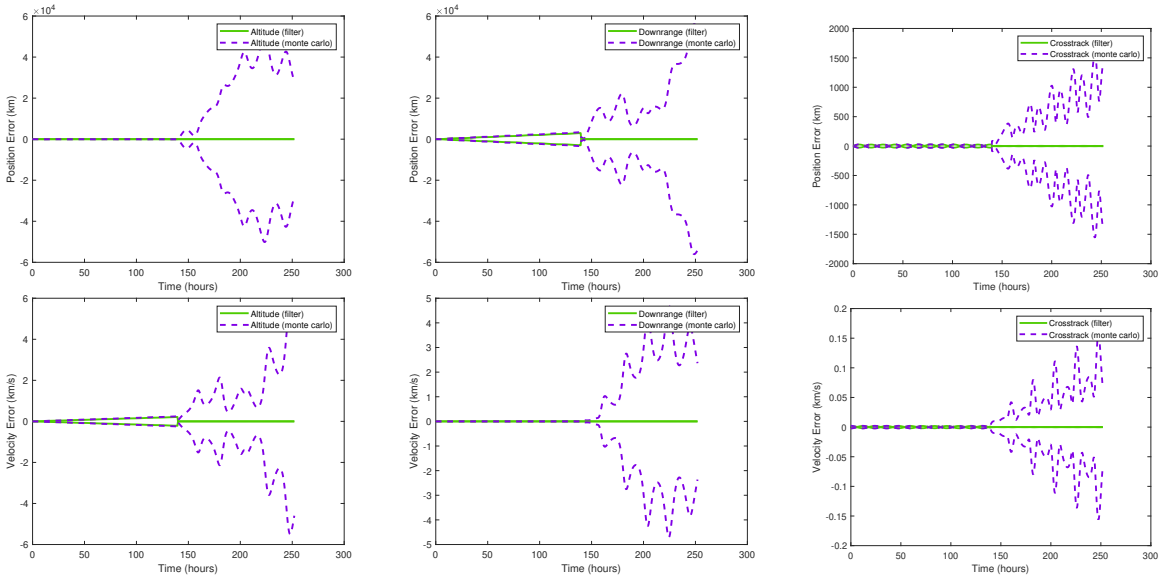


Fig. F.48: GMM Measurement at 140 Hours, All Observations

F.5 EXTENDED STEP-BACK KALMAN FILTER

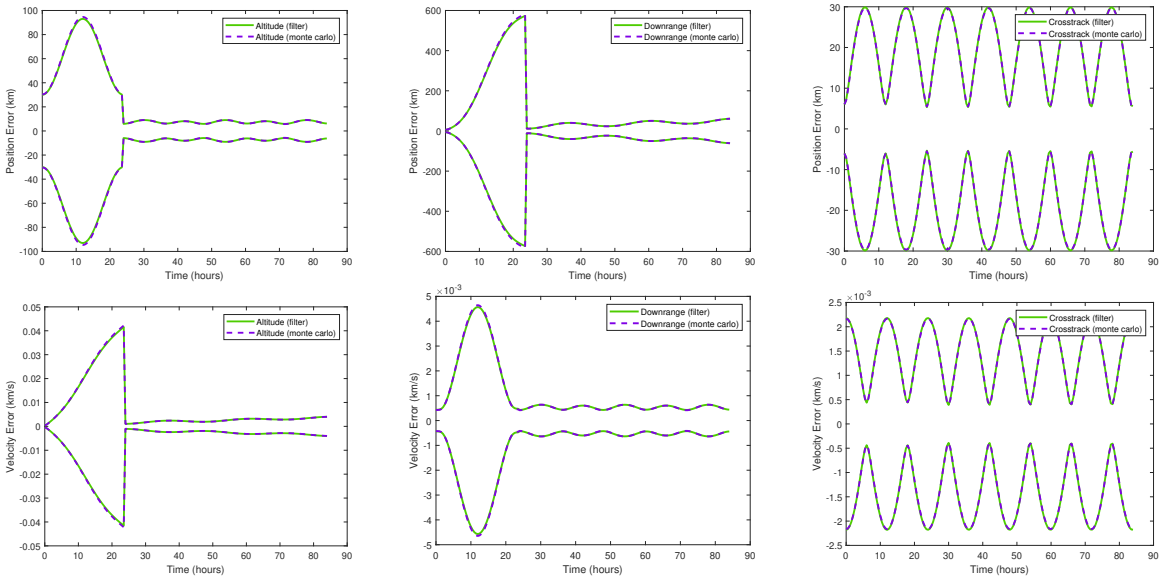


Fig. F.49: ESBKF Measurement at 24 Hours, OBS 1 Only

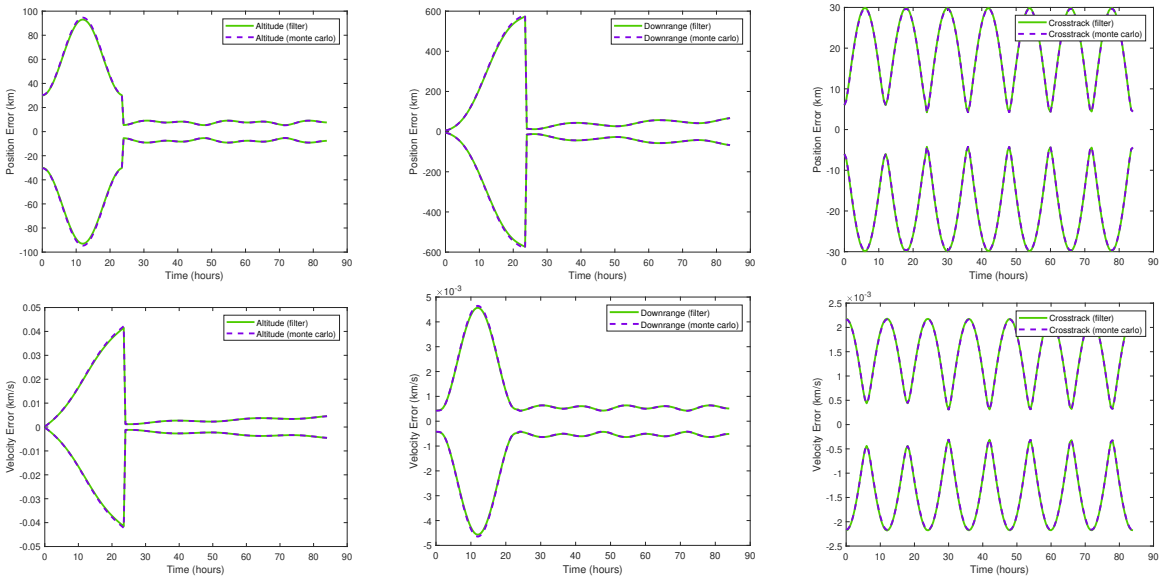


Fig. F.50: ESBKF Measurement at 24 Hours, OBS 2 Only

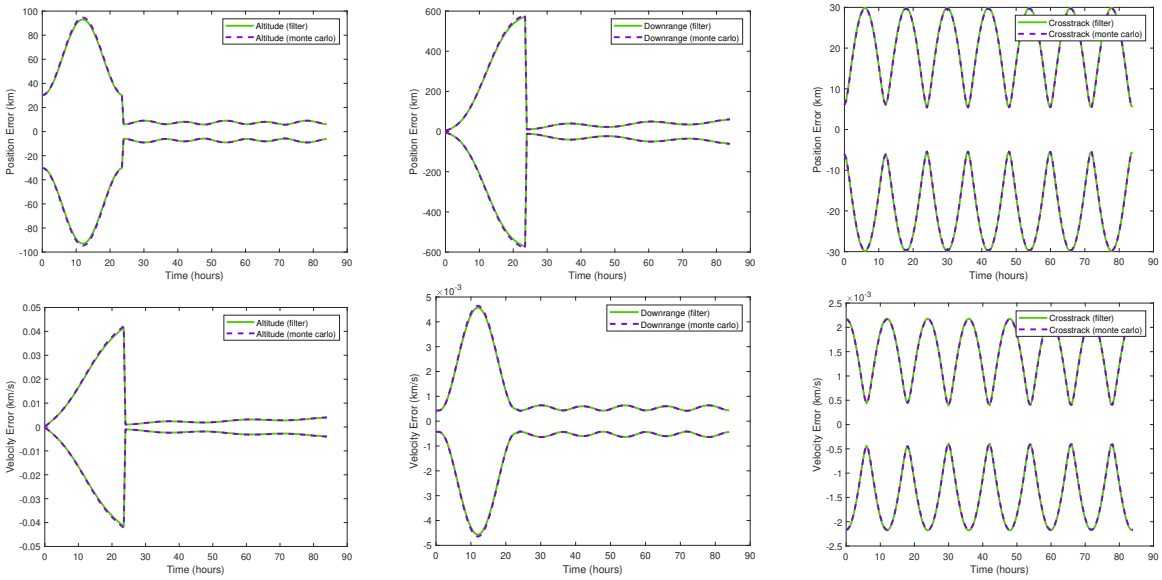


Fig. F.51: ESBKF Measurement at 24 Hours, OBS 3 Only

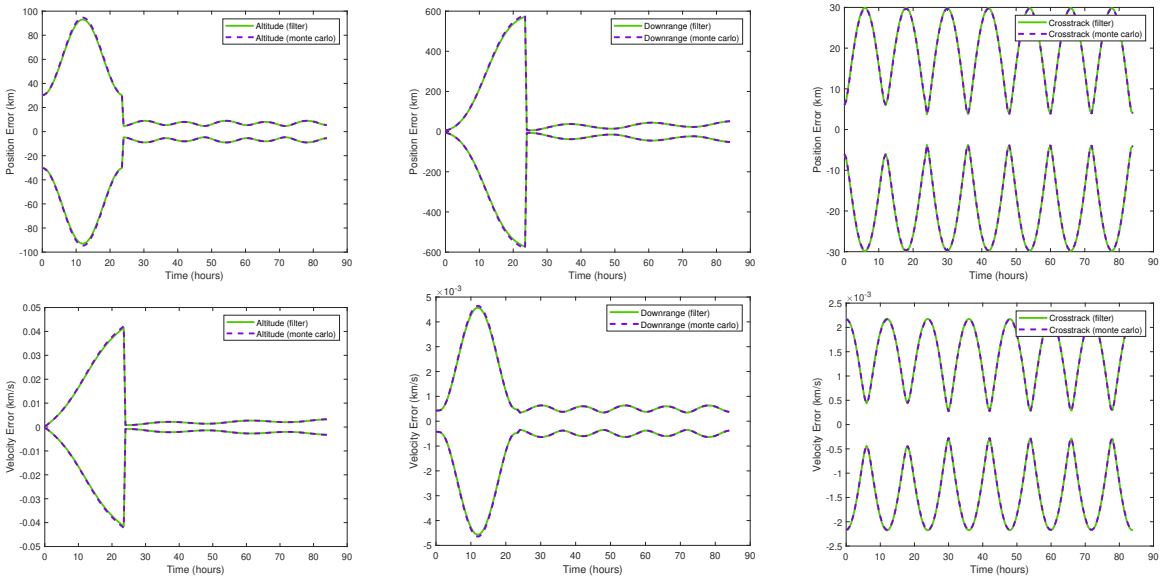


Fig. F.52: ESBKF Measurement at 24 Hours, All Observations

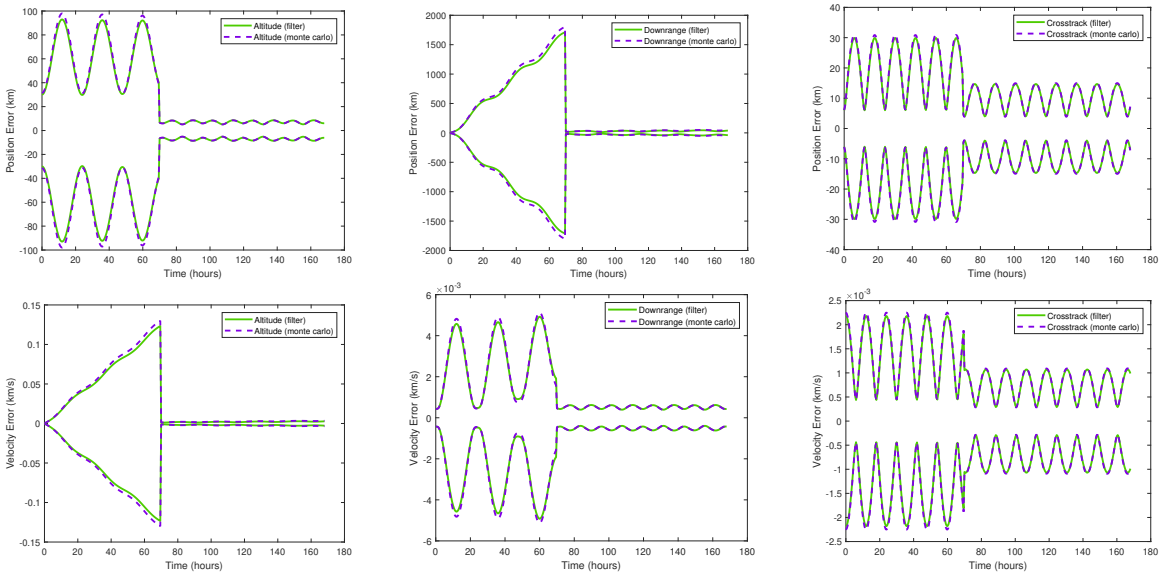


Fig. F.53: ESBKF Measurement at 70 Hours, OBS 1 Only

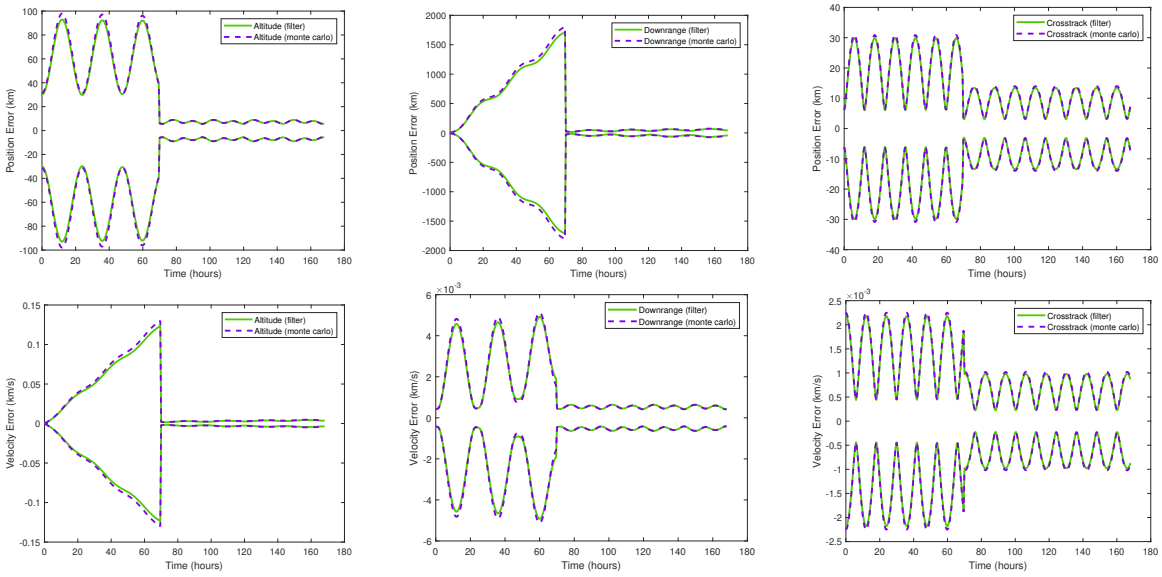


Fig. F.54: ESBKF Measurement at 70 Hours, OBS 2 Only

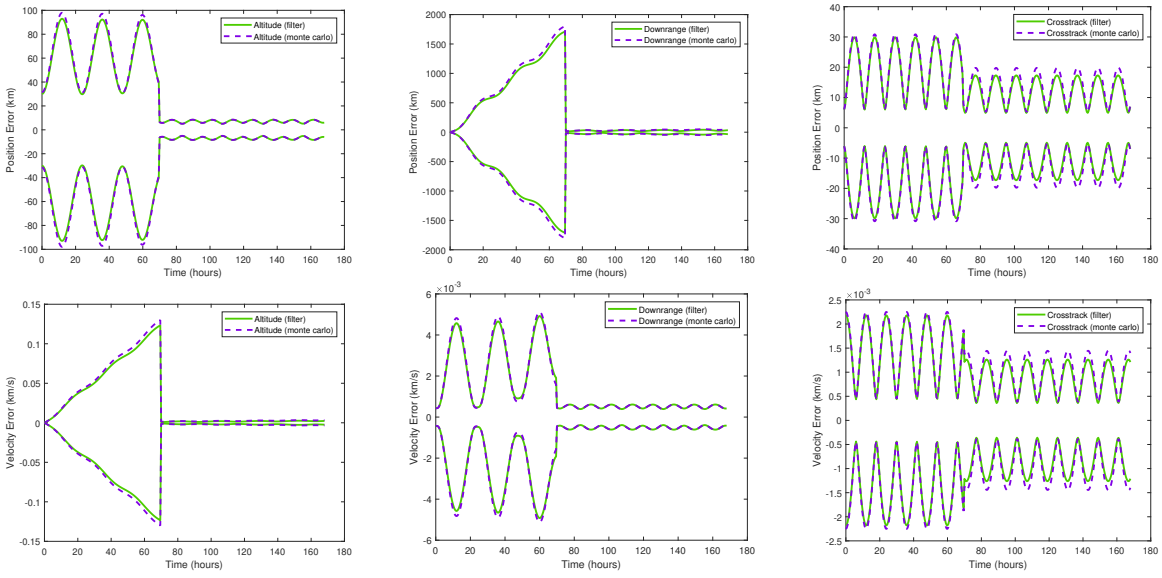


Fig. F.55: ESBKF Measurement at 70 Hours, OBS 3 Only

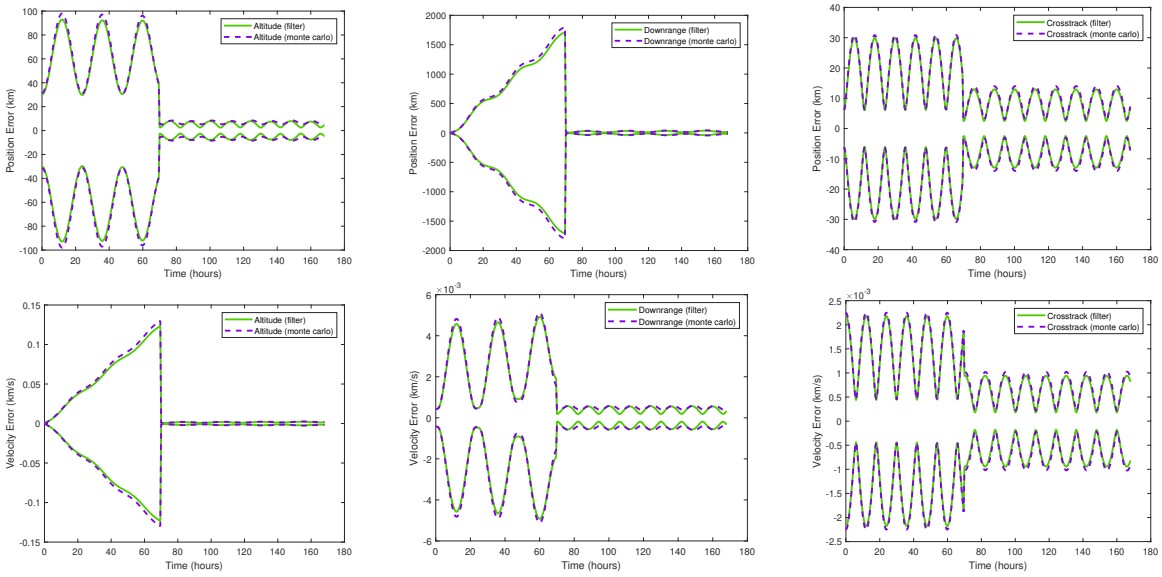


Fig. F.56: ESBKF Measurement at 70 Hours, All Observations

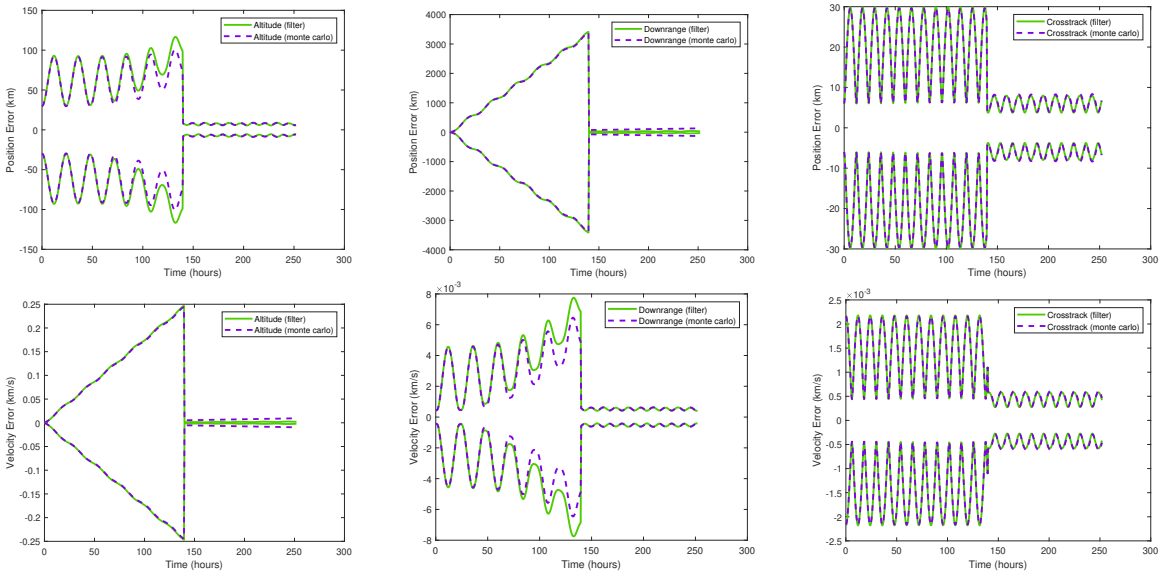


Fig. F.57: ESBKF Measurement at 140 Hours, OBS 1 Only

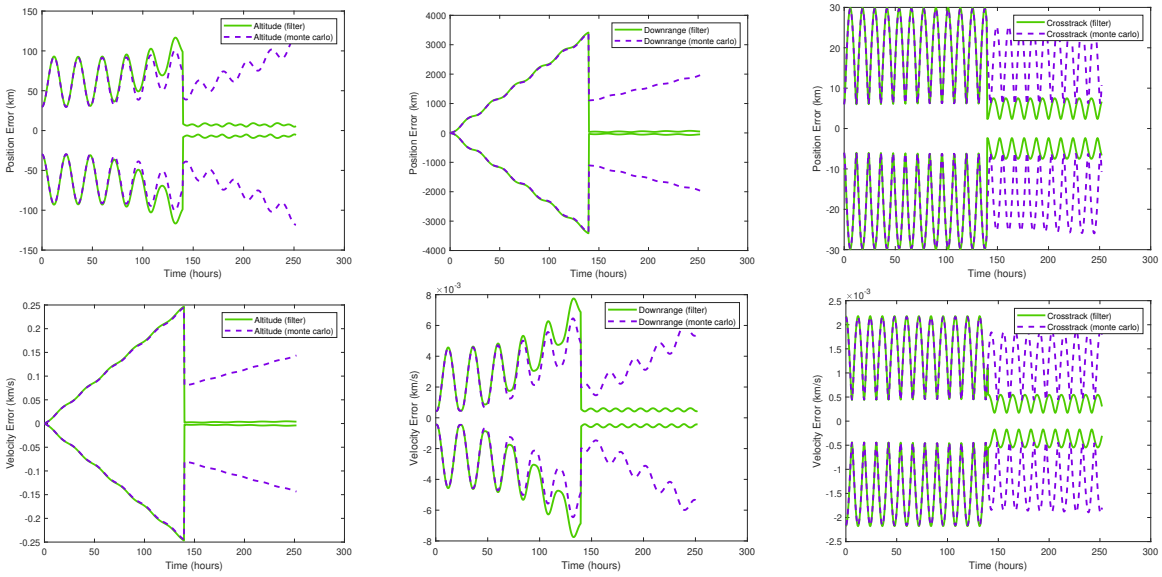


Fig. F.58: ESBKF Measurement at 140 Hours, OBS 2 Only

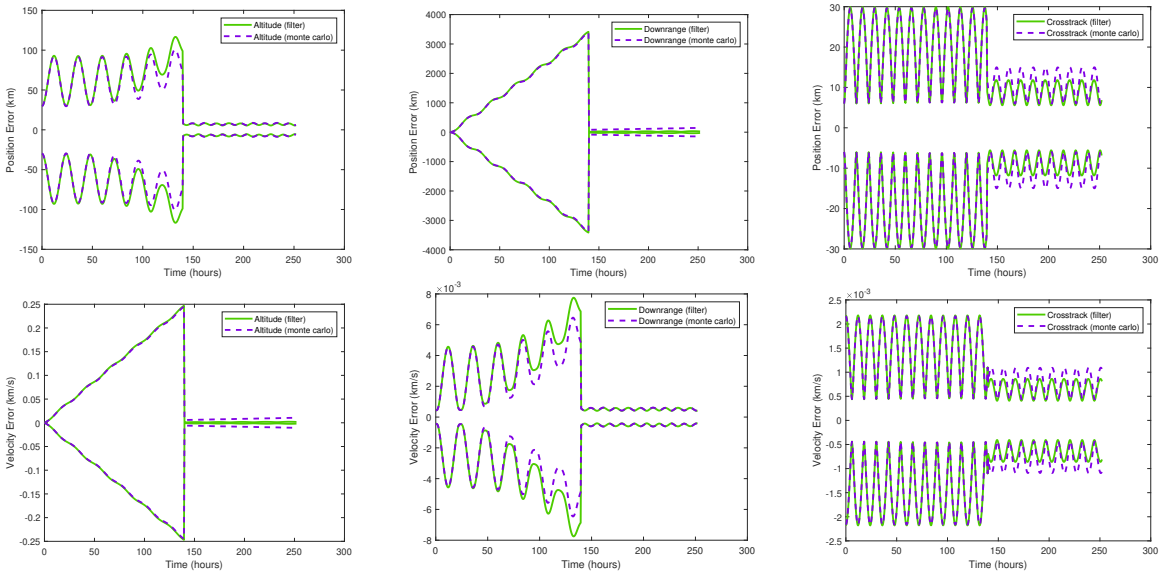


Fig. F.59: ESBKF Measurement at 140 Hours, OBS 3 Only

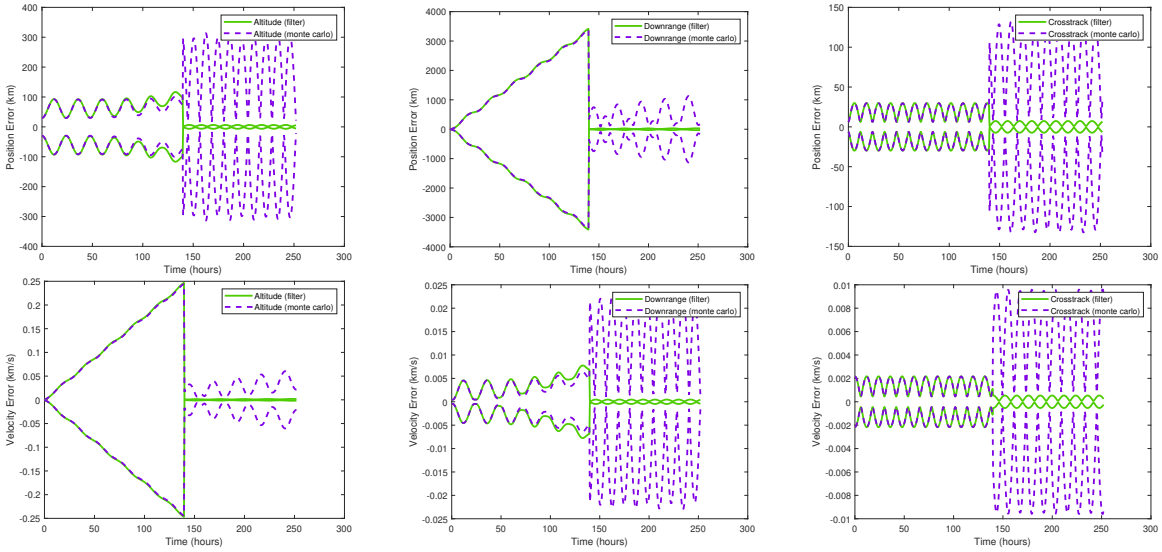


Fig. F.60: ESBKF Measurement at 140 Hours, All Observations

CURRICULUM VITAE

Louis M Tonc

Education

- M.S. Aerospace Engineering, Utah State University, Logan, Utah, 2016.
- B.S. Mechanical Engineering, Utah State University, Logan, Utah, 2010.

Work Experience

- 2011 - 2013 Associate Mine Engineer, Pacific Minerals Inc. - Bridger Coal Company, Point of Rocks, WY.

  Oversaw planning and construction for mining supplies delivery road. Assisted in major ventilation study for underground mine. Oversaw multiple projects for the underground mine on the surface.