

Nonlinear Bayesian Filtering Based on Mixture of Orthogonal Expansions

Syed Amer Ahsan Gilani

Submitted for the Degree of
Doctor of Philosophy
from the
University of Surrey



Surrey Space Centre
Faculty of Electronics and Physical Sciences
University of Surrey
Guildford, Surrey GU2 7XH, UK.

Mar 2012

© Syed Amer Ahsan Gilani 2012

I begin with the name of ALMIGHTY ALLAH (GOD)
who is most Merciful and most Beneficent (Al-Quran)

Abstract

This dissertation addresses the problem of parameter and state estimation of nonlinear dynamical systems and its applications for satellites in Low Earth Orbits. The main focus in Bayesian filtering methods is to recursively estimate the state *a posteriori* probability density function conditioned on available measurements. Exact optimal solution to the nonlinear Bayesian filtering problem is intractable as it requires knowledge of infinite number of parameters. Bayes' probability distribution can be approximated by mixture of orthogonal expansion of probability density function in terms of higher order moments of the distribution. In general, better series approximations to Bayes' distribution can be achieved using higher order moment terms. However, use of such density function increases computational complexity especially for multivariate systems.

Mixture of orthogonally expanded probability density functions based on lower order moment terms is suggested to approximate the Bayes' probability density function. The main novelty of this thesis is development of new Bayes' filtering algorithms based on single and mixture series using a Monte Carlo simulation approach. Furthermore, based on an earlier work by Culver [1] for an exact solution to Bayesian filtering based on Taylor series and third order orthogonal expansion of probability density function, a new filtering algorithm utilizing a mixture of orthogonal expansion for such density function is derived. In this new extension, methods to compute parameters of such finite mixture distributions are developed for optimal filtering performance. The results have shown better performances over other filtering methods such as Extended Kalman Filter and Particle Filter under sparse measurement availability. For qualitative and quantitative performance the filters have been simulated for orbit determination of a satellite through radar measurements / Global Positioning System and optical navigation for a lunar orbiter. This provides a new unified view on use of orthogonally expanded probability density functions for nonlinear Bayesian filtering based on Taylor series and Monte Carlo simulations under sparse measurements.

Another new contribution of this work is analysis on impact of process noise in mathematical models of nonlinear dynamical systems. Analytical solutions for nonlinear differential equations of motion have a different level of time varying process noise. Analysis of the process noise for Low Earth Orbital models is carried out using the Gauss Legendre Differential Correction method. Furthermore, a new parameter estimation algorithm for Epicyclic orbits by Hashida and Palmer [2], based on linear least squares has been developed.

The foremost contribution of this thesis is the concept of nonlinear Bayesian estimation based on mixture of orthogonal expansions to improve estimation accuracy under sparse measurements.

Acknowledgements

Working towards completion of this project has been the most challenging part of my life. First of all I am thankful to Almighty ALLAH (GOD) who gave me the opportunity and skill to undertake this project. Then I am thankful to my supervisor Dr P L Palmer who generously helped me and taught me very well. Next I am thankful to my sponsors at National University of Sciences and Technology (NUST) Pakistan and colleagues at Surrey Space Centre which includes David Wokes, Kristian Kristiansen, Luke Sauter, Andrew Auman, Chris Bridges and Naveed Ahmed. And last but not the least my wife and son Ali who endured this journey together patiently yet cheerfully and rest of the family members in Pakistan who always wished me very well.

Table of Contents

Abstract.....	3
Table of Contents.....	5
List of Figures	9
List of Acronyms.....	13
1 Introduction.....	18
1.1 Overview	18
1.2 Motivation	18
1.3 Discussion of Problem	21
1.4 Aims and Objectives	24
1.4.1 Aims	24
1.4.2 Objectives	24
1.5 Structure of Thesis	24
1.6 Novelty	25
1.7 Publications	26
2 Literature Survey	27
2.1 Nonlinear Bayesian Recursive Filtering.....	27
2.1.1 Gaussian Based Methods.....	27
2.1.2 Gaussian Mixture Model Based Methods.....	28
2.1.3 Sequential Monte Carlo Methods.....	29
2.1.4 Orthogonal Expansion Based Methods	30
2.1.5 Numerical Based Methods.....	31
2.1.6 Variational Bayesian Methods	31
2.2 Parameter Estimation	32
2.3 Satellite Orbital Dynamics.....	32
2.4 Satellite Relative Motion.....	33
2.5 Summary	34
3 Analysis of Fidelities of Linearized Orbital Models	35
3.1 Introduction	35
3.2 Methodology for Fitting Approximate Models to Nonlinear Data.....	37
3.3 Two Body Equation Review	40
3.3.1 Kepler's Equation	42
3.3.2 Conversion from Perifocal to ECI Coordinates	43
3.4 Perturbation Due to Oblate Earth – J_2	45

3.5	Analysis of Absolute Satellite Orbital Dynamics.....	50
3.5.1	Analysis of Kepler's Equation	51
3.5.1.1	Unperturbed Two Body Equation	51
3.5.1.2	J_2 Perturbed Two Body Equation	54
3.5.2	Epicyclic Motion of Satellite about an Oblate Planet.....	57
3.5.3	Conclusion	64
3.6	Relative Motion between Satellites	64
3.7	Analysis of Relative Motion	66
3.7.1	Hill Clohessy Wiltshire Model.....	67
3.7.2	Orbit Eccentricity	76
3.7.3	Semi Major Axis and Inclination	76
3.7.4	J_2 Modified HCW Equations by Schweighart and Sedwick	79
3.7.5	Conclusion	83
3.8	Free Propagation Error Growth	84
3.9	Summary	86
4	Epicycle Orbit Parameter Filter.....	87
4.1	Introduction	87
4.2	Secular Variations in Epicycle Orbital Coordinates	91
4.3	Development of an Epicycle Parameter Filter	93
4.3.1	Reference Nonlinear Satellite Trajectory.....	93
4.3.2	Least Squares Formulation	94
4.3.3	Determination of Semi Major Axis " a " and Inclination " I_0 "	96
4.3.4	Determination of " ζ_P " and " η_P "	97
4.4	Parameter Estimation Accuracy	100
4.5	Error Statistics in Orbital Coordinates at Different I_0	103
4.6	Time History of Errors in Epicycle Coordinates.....	105
4.7	Time History of Errors in Epicycle Coordinates Without Estimation	108
4.8	Free Propagation Secular Error Growth	110
4.9	Summary	113
5	Development of Gram Charlier Series and its Mixture Particle Filters	114
5.1	Introduction	114
5.2	Fundamentals of Particle Filters	118
5.2.1	Monte Carlo Integration	118
5.2.2	Bayesian Importance Sampling	119
5.2.3	Sequential Importance Sampling	120
5.2.4	Degeneration of Particles and its Minimization.....	122
5.2.5	Generic Bootstrap Particle Filter Algorithm.....	124

5.2.6 Parametric Bootstrap Particle Filtering Algorithms	124
5.2.6.1 Gaussian Particle Filter	124
5.2.6.2 Gaussian Sum Particle Filter	125
5.3 Gram Charlier Series	127
5.3.1 Univariate GCS	127
5.3.2 Multivariate GCS	128
5.4 Gram Charlier Series Mixture Model	129
5.4.1 Univariate Gram Charlier Series Mixture Model	130
5.4.2 Multivariate GCSMM	132
5.5 Random Number Generation	136
5.5.1 GCS Random Number Generator using Acceptance Rejection	136
5.5.2 Gram Charlier Series Random Number Generator using Gaussian Copula ...	141
5.6 Gram Charlier Series and its Mixture Particle Filtering	142
5.6.1 Single Gram Charlier Series Particle Filtering	143
5.6.2 Gram Charlier Series Mixture Particle Filtering	148
5.7 Experiments – Nonlinear Simple Pendulum	150
5.7.1 Atmospheric Drag	150
5.7.2 Wind Gust	157
5.7.3 Experiment – Radar Based Orbit Determination	160
5.8 Summary	175
6 Development of Mixture Culver Filter	177
6.1 Introduction	177
6.2 Continuous Discrete Nonlinear Filtering Problem	180
6.3 Culver Filter	182
6.4 Mixture Culver Filter	185
6.4.1 Time Update	186
6.4.2 Measurement Update	191
6.5 Orbit Determination using Radar Measurements	196
6.5.1 State Uncertainty and Sparse Measurements	197
6.5.2 Discussion	205
6.6 Lunar Orbital Navigation	206
6.7 Summary	210
7 Conclusion and Future Work	211
7.1 Introduction	211
7.2 Concluding Summary	211
7.3 Research Achievements	212
7.4 Extensions and Future Work	213

References.....	215
Appendix A: Transformation Routines	223
Appendix B: Partial for State Transition Matrix Kepler's Equation	225
Appendix C: Epicycle Coefficients for Geopotential Zonal Harmonic Terms up to J_4	227
Appendix D: Partial for Epicyclic Orbit Analysis.....	230
Appendix E: Analytical Solution of Modified HCW Equations by SS	235
Appendix F: Partial for Modified HCW Equations by SS.....	237
Appendix G: Analytical Solution of Integrals for GCSMM Time Update.....	243

List of Figures

Figure 1-1:	Block description of state estimation.....	19
Figure 1-2:	Block description of Bayesian prediction and update stages.	22
Figure 3-1:	The concept of divergence.	36
Figure 3-2:	Concept of methodology for linearized orbital analysis..	40
Figure 3-3:	Earth Central Inertial (ECI) Coordinate frame.....	41
Figure 3-4:	Orbital geometry for Kepler's equation.	43
Figure 3-5:	Geometrical description of geocentric latitude ϕ and longitude Λ	45
Figure 3-6:	Time history of a satellite orbit in ECI coordinates	47
Figure 3-7:	Time history of a satellite orbit in ECI coordinates.....	47
Figure 3-8:	Time history of variations (Δ) in orbital elements.....	48
Figure 3-9:	Time history of variations (Δ) in orbital elements.....	48
Figure 3-10:	Time history of variations (Δ) in angular quantities of orbital elements	49
Figure 3-11:	Time history of variations (Δ) in angular quantities of orbital elements	49
Figure 3-12:	Illustration of the Local Vertical Local Horizontal (LVLH) system	50
Figure 3-13:	Time history of position errors for analytic solution of Kepler's equation.	53
Figure 3-14:	Time history of velocity errors for analytic solution of Kepler's equation	53
Figure 3-15:	Time history of position errors for analytic solution of Kepler equation.....	55
Figure 3-16:	Time history of velocity errors for analytic solution of Kepler's equation.	55
Figure 3-17:	Time history of position errors for analytic solution of Kepler's equation.	56
Figure 3-18:	Time history of velocity errors for analytic solution of Kepler's equation.	56
Figure 3-19:	Geometrical representation of epicycle coordinates	58
Figure 3-20:	Time history of position errors for epicycle orbit.....	61
Figure 3-21:	Time history of velocity errors for epicycle orbit.....	62
Figure 3-22:	Time history of position errors for epicycle orbit.....	63
Figure 3-23:	Time history of velocity errors for epicycle orbit.....	63
Figure 3-24:	Illustration of the satellite relative motion coordinate system.....	65
Figure 3-25:	Geometry of the free orbit ellipse for relative motion	67
Figure 3-26:	Illustration of " <i>free orbit ellipse</i> " relative orbit.....	72
Figure 3-27:	Time history of position errors for HCW equations	73
Figure 3-28:	Time history of velocity errors HCW equations.....	74
Figure 3-29:	Time history of position errors HCW equations	75
Figure 3-30:	Time history of velocity errors for HCW equations	76
Figure 3-31:	Maximum position errors for HCW equations	77

Figure 3-32:	Maximum velocity errors for HCW equations	77
Figure 3-33:	Maximum position errors (radial direction) for HCW model	78
Figure 3-34:	Maximum position errors (in-track direction) for HCW model	78
Figure 3-35:	Maximum position errors (cross-track direction) for HCW model	79
Figure 3-36:	Time history of position errors for SS model after using optimal initial conditions. ...	81
Figure 3-37:	Time history of velocity errors for SS model after using optimal initial conditions. ...	82
Figure 3-38:	Time history of position errors for SS model without modifying initial conditions.....	82
Figure 3-39:	Time history of velocity errors for SS model without modifying initial conditions.....	83
Figure 3-40:	Time history of growth of position errors for HCW model	84
Figure 3-41:	Time history of growth of position errors for SS model	85
Figure 3-42:	Time history of growth of position errors for epicycle model.....	85
Figure 4-1:	The plot depicts the dominant linear secular growth.....	92
Figure 4-2:	The plot depicts the dominant linear secular growth.....	92
Figure 4-3:	Flow chart of the Epicycle Parameter Filter (EPF)	99
Figure 4-4:	J_2 epicycle coefficients for radial offset (ρ_2), and secular drift (κ_2, ϑ_2)	100
Figure 4-5:	J_2 epicycle coefficients for the radial offset (ρ_2), and secular drift (κ_2, ϑ_2)	100
Figure 4-6:	Percentage estimation errors (Δ)	102
Figure 4-7:	Estimation errors (Δ) for inclination.....	103
Figure 4-8:	Maximum absolute errors.	104
Figure 4-9:	Maximum absolute errors	104
Figure 4-10:	Maximum absolute errors	105
Figure 4-11:	Time history of errors (Δ).....	106
Figure 4-12:	Time history of errors (Δ)	106
Figure 4-13:	Time history of errors (Δ)	107
Figure 4-14:	Time history of errors (Δ)	107
Figure 4-15:	Time history of errors (Δ).....	107
Figure 4-16:	Time history of errors (Δ)	108
Figure 4-17:	Time history of errors (Δ)	108
Figure 4-18:	Time history of errors (Δ).....	109
Figure 4-19:	Time history of errors (Δ)	109
Figure 4-20:	Time history of errors (Δ).	109
Figure 4-21:	Time history of errors (Δ).	110
Figure 4-22:	Time history of errors (Δ)	110
Figure 4-23:	Time history of radial coordinate	111
Figure 4-24:	Time history of errors (Δ)	111
Figure 4-25:	Time history of errors (Δ)	112
Figure 4-26:	Time history of errors (Δ).	112

Figure 5-1:	Discrete filtering.....	115
Figure 5-2:	Block description of Bayesian prediction and update stages	116
Figure 5-3:	SIR.....	123
Figure 5-4:	The comparison of true exponential PDF	131
Figure 5-5:	The comparison of true uniform PDF.....	132
Figure 5-6:	Gaussian kernel based non-parametric density estimation.	138
Figure 5-7:	Single Gaussian PDF contours.....	139
Figure 5-8:	Single GCS (5 th order) PDF contours	139
Figure 5-9:	Three components GMM PDF contours.....	140
Figure 5-10:	Three components GCSMM (5 th order) PDF	140
Figure 5-11:	Three components GCSMM (3 rd order) PDF.....	141
Figure 5-12:	Comparison of time history of errors in angular position	155
Figure 5-13:	Comparison of time history of errors angular velocity.	156
Figure 5-14:	Comparison of time history of errors in angular position	158
Figure 5-15:	Comparison of time history of errors angular velocity	159
Figure 5-16:	Measurement model description in Topocentric Coordinate System.	161
Figure 5-17:	Time history of errors (Δ) in ECI X (top), \dot{X} (middle), and Y (bottom). The measurement frequency is 0.2 Hz.....	165
Figure 5-18:	Time history of errors (Δ) in ECI \dot{Y} (top), Z (middle), and \dot{Z} (bottom). The measurement frequency is 0.2 Hz.....	166
Figure 5-19:	Time history of magnitude of errors in position $ \Delta r $ m (top) and velocity $ \Delta v $ m/s (bottom). Measurement frequency is 0.2 Hz.	167
Figure 5-20:	Time history of errors (Δ) in ECI X (m) and \dot{X} (m/s) after one orbital period T , where $T = \sim 97$ min.	168
Figure 5-21:	Time history of errors (Δ) in ECI Y (top), \dot{Y} (middle), and Z (bottom). The measurement frequency is 0.2 Hz after one orbital period T , where $T = \sim 97$ min.	169
Figure 5-22:	Time history of errors (Δ) in ECI \dot{Z} (top), $ \Delta r $ m (middle), and $ \Delta v $ m/s (bottom). The measurement frequency is 0.2 Hz after one orbital period T , where $T = \sim 97$ min.	170
Figure 5-23:	Time history of position errors (Δ) in ECI coordinates for a GSPF.	172
Figure 5-24:	Time history of positional covariance for a GSPF.....	172
Figure 5-25:	Time history of position errors (Δ) in ECI coordinates for a GCSMPF.	173
Figure 5-26:	Time history of positional covariance for a GCSMPF.....	173
Figure 5-27:	Time history of ECI position errors Δ for GCSMPF during subsequent orbital periods, (a) 2 nd orbital period, (b) 3 rd orbital period, where $T = 5805$ sec.....	174
Figure 5-28:	Time history of ECI position errors Δ for GCSMPF during subsequent orbital periods, (a) 4 th orbital period, (b) 5 th orbital period, where $T = 5805$ sec.....	174

Figure 5-29: Time history of ECI position errors Δ for GCSMPF during subsequent orbital periods, (a) 6 th orbital period, (b) 7 th orbital period, where $T = 5805$ sec.....	175
Figure 6-1: Continuous-discrete filtering	177
Figure 6-2: The block description of continuous-discrete filtering.	178
Figure 6-3: Time history of absolute position errors Δ in ECI coordinates	198
Figure 6-4: Time history of absolute velocity errors Δ in ECI coordinates	198
Figure 6-5: Time history of absolute errors Δ in ECI coordinates.....	199
Figure 6-6: Time history of absolute errors Δ in ECI coordinates.....	201
Figure 6-7: Time history of absolute RMSE in ECI X_I	201
Figure 6-8: Time history of absolute RMSE in ECI Y_I	202
Figure 6-9: Time history of absolute RMSE in ECI Z_I	202
Figure 6-10: Time history of absolute position errors Δ in ECI coordinates	203
Figure 6-11: Time history of RMSE in ECI coordinates (X-axis) for filters.....	203
Figure 6-12: Time history of RMSE in ECI coordinates (Y-axis) for filters.....	204
Figure 6-13: Time history of RMSE in ECI coordinates (Z-axis) for filters.....	204
Figure 6-14: Lunar navigation system description	207
Figure 6-15: Time history of absolute position errors (Δ) in Cartesian positions for Culver framework under sparse measurements.	209
Figure 6-16: Time history of absolute velocity errors (Δ) in Cartesian velocities for Culver framework under sparse measurements.	209

List of Acronyms

AOCS – Attitude and Orbit Control Systems

AR – Acceptance Rejection

AFB – Air force Base

CF – Culver Filter

CKE – Chapman Kolmogorov Equation

DSSM – Discrete State Space Model

ECI – Earth Central Inertial

ECEF – Earth Central Earth Fixed

EKF – Extended Kalman Filter

EPF – Epicycle Parameter Filter

FPKE – Fokker Planck Kolmogorov Equation

FD – Finite Difference

GPS – Global Positioning System

GMM – Gaussian Mixture Model

GCSMM – Gram Charlier Series Mixture Model

GCS – Gram Charlier Series

GSF – Gaussian Sum Filter

GPF – Gaussian Particle Filter

GLDC – Gauss Legendre Differential Correction

GCSPF – Gram Charlier Series Particle Filter

GCSMPF - Gram Charlier Series Mixture Particle Filter

GBF – Grid based Filters

HCW – Hill Clohessy Wiltshire

IC – Initial Condition

ISE – Integrated Square Error

KF – Kalman Filter

LEO – Low Earth Orbits

LVLH – Local Vertical Local Horizontal

MC – Monte Carlo

MCF – Mixture Culver Filter

MMSE – Minimum Mean Square Error

MAP – *maximum a posteriori*

MLE – maximum likelihood estimates

NORAD – North American Aerospace Defence Command

OD – Orbit Determination

OBC – Onboard Computer

PDF – Probability Density Function

PF – Particle Filter

RBPF – Rao-Blackwell Particle Filter

RAAN – Right Ascension of the Ascending Node

SIR – Sampling Importance Resampling

SS – Schweighart and Sedwick

SSC – Surrey Space Centre

SDE – Stochastic Differential Equation

SAR – Synthetic Aperture Radar

SMC – Sequential Monte Carlo

SIS – Sequential Importance Sampling

TLE – Two Line Element

VLSI – Very Large Scale Integrated circuits

List of Symbols

\mathbf{x}_0	IC or parameters of dynamical system
\mathbf{x}_k	True state of a dynamic system at k^{th} instant of time
$\hat{\mathbf{x}}_k$	Estimated state of a dynamic system at k^{th} instant of time
\mathbf{P}_k	Covariance matrix at k^{th} instant of time
$\mathbf{P}_k^{(3)}$	Coskewness tensor at k^{th} instant of time
$\mathbf{P}_k^{(4)}$	Cokurtosis or fourth order tensor at k^{th} instant of time
$\mathbf{P}_k^{(5)}$	Fifth order tensor at k^{th} instant of time
$\kappa_{(.)}$	Cumulants of PDF
$[\vec{r}, \vec{v}]$	Position and velocity vectors in ECI coordinate system
\mathbf{Z}	Nonlinear trajectory in ECI coordinate system
\mathcal{P}	Analytical trajectory in ECI coordinate system
\mathbf{q}	Process noise
$E[.]$	Expectation operator
$E(t)$	Eccentric anomaly
\mathbf{F}	Jacobian matrix
μ_E	Gravitational parameter of Earth
$[X, Y, Z]$	ECI position coordinate system
$[\dot{X}, \dot{Y}, \dot{Z}]$	ECI velocity coordinate system
G_E	Earth's gravitational constant
M_E	Mass of Earth
Φ	Gravitational potential function for spherical Earth
Ω	RAAN
ω	Argument of perigee
ν	True anomaly
\vec{e}	Eccentricity vector

a	Semi major axis
ε	Orbital energy
n	Mean motion
M	Mean anomaly
t_p	Time of perigee passage
t_E	Time of equator crossing
$[\hat{X}, \hat{Y}]$	Orbital coordinates of a satellite in Perifocal coordinate system
$\mathbf{P}, \mathbf{Q}, \mathbf{W}$	Vectors to define Perifocal coordinate system
$\mathfrak{R}(\cdot)$	Rotation matrix
R_E	Radius of Earth
\mathfrak{U}	Gravitational potential function for non spherical Earth
\mathcal{U}	Uniform random number
$P_l[\cdot]$	Legendre polynomial of degree “ l ”
J_l	Coefficient of zonal spherical harmonic representing shape of Earth
Λ	Geocentric longitude of Earth
$\hat{I}, \hat{J}, \hat{K}$	Vectors to define ECEF coordinate system
\mathbf{a}_G	Perturbation acceleration due to zonal gravitational harmonics
\mathbf{a}_D	Perturbation acceleration due to atmospheric drag
x, y, z	Vectors to define LVLH coordinate system
I	Instantaneous inclination of orbital plane for Epicycle orbit
λ	Instantaneous argument of latitude for Epicycle orbit
v_r	Instantaneous radial velocity for Epicycle orbit
v_θ	Instantaneous azimuthal velocity for Epicycle orbit
$[\xi_P, \eta_P]$	Non singular Epicycle parameters
A	Epicycle or relative orbit amplitude
$p(\mathbf{x}_k \mathbf{y}_{1:k})$	Bayes’ <i>a posteriori</i> PDF
$\delta(\cdot)$	Dirac delta function

$p(\mathbf{x}_k \mathbf{y}_{1:k})$	Proposal PDF
$w_k^{(i)}$	Weight of i^{th} particle at k^{th} instant of time
$\mathbf{x}_k^{(i)}$	i^{th} particle at k^{th} instant of time
$p_g(\cdot)$	Gaussian PDF
$\mathcal{N}(\cdot)$	Gaussian PDF (alternate symbol)
$p_{gmm}(\cdot)$	GMM PDF
$p_{gcs}(\cdot)$	GCS PDF
$p_{gcsm}(\cdot)$	GCSMM PDF
\otimes	Kronecker product
C_D	Coefficient of drag
\mathbf{W}	Continuous time white Gaussian noise
\mathbf{w}_k	Discrete time white Gaussian noise
β	Brownian motion
ρ	Radar site to satellite position vector
\mathbf{R}_S	ECI coordinates of radar site

Note: Any reuse of symbols is defined appropriately within the text.

1 Introduction

1.1 Overview

A dynamical system is described by a mathematical model either in discrete time or continuous time. In discrete time the evolution is considered at fixed discrete instants usually with positive integer numbers, whereas, in continuous time the progression of time is smooth occurring at each real number. No mathematical model is perfect. There are sources of uncertainty in any mathematical model of a system due to approximations of physical effects. Moreover, these models do not account for system dynamics driven by disturbances which can neither be controlled nor modelled deterministically. For example, if a pilot wants to steer an aircraft at a certain angular orientation, the true response will be different due to wind buffeting, imprecise actuator response and inability to accurately generate the desired response from hands on the control stick [3]. These uncertainties can be approximated as noise in the system dynamics. The numerical description of current configuration of a dynamical system is called a state [4]. For a particular dynamical system one needs to obtain knowledge of the possible motion or state of the system. The state is usually observed indirectly by sensors which provide output data signals described as a function of state. Sensors do not provide perfect and complete data about the system as they introduce their own system dynamics and distortions [3]. Moreover, the measurements are always corrupted by noise.

Estimation of state can be understood as the process of acquiring knowledge about possible motions of a particular dynamic system. It utilizes prior information for prediction of the estimated state, extracts noisy measurements and characterizes dynamic system uncertainties. Figure 1-1 explains block methodology of the estimation process. The true dynamic system and measurement devices can be considered as a physical (hardware) layer of the complete process. The mathematical model of the dynamic system and measurement model along with their noise characterization and prior state information is used by estimation algorithm to provide current state estimates and associated uncertainties. This could be understood as software layer. Bayes' formula describes how Probability Density Function (PDF) or *belief in predicted state* of a dynamic system is modified based on evidence from the measurement data *the likelihood function of state* [5].

1.2 Motivation

Most of the dynamical systems in the real world are nonlinear. This intrigues researchers and scientists to study more about their characteristics and behaviour. In the context of state estimation for

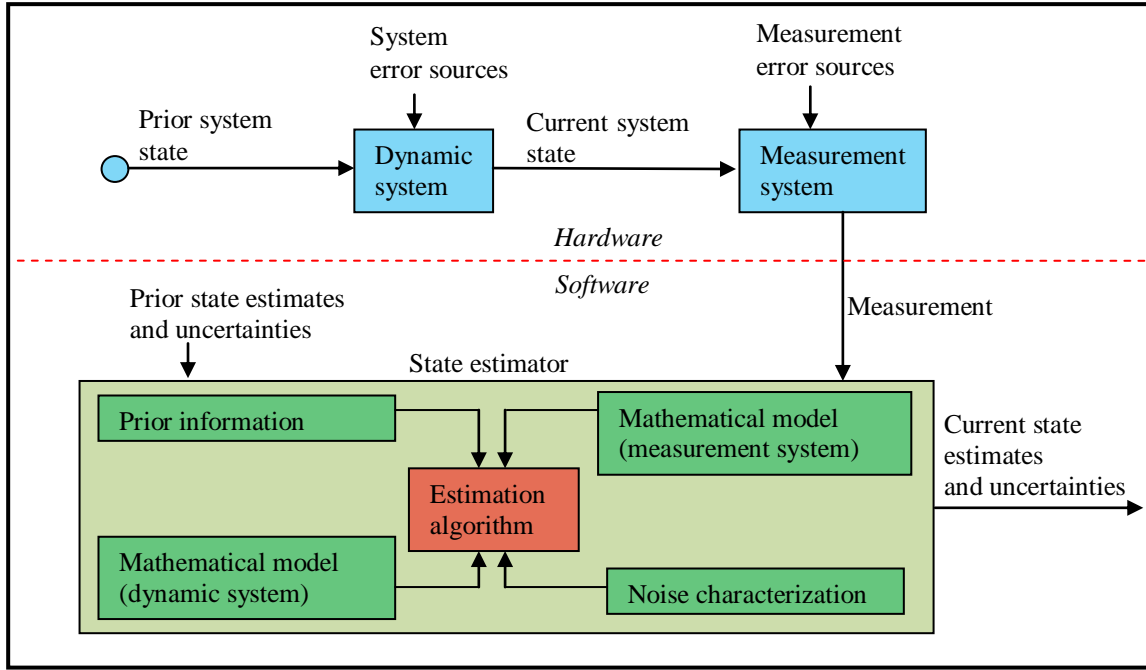


Figure 1-1: Block description of state estimation. The current true state of a system is measured and provided to state estimator (hardware layer). State estimator utilizes mathematical models, prior state information, characterization of noise and estimation algorithm to obtain current state estimates and uncertainties (software layer).

nonlinear dynamical systems, knowledge about time evolution of their PDF is very crucial. The form of this PDF is complicated and it is difficult to describe it with some tractable function. In general this density function cannot be characterized by a finite set of parameters e.g., moments unlike linear systems where full description up to second order statistics is sufficient [6]. Therefore, linear systems are sometimes referred as Gaussian based systems, owing to their complete description by first two moments. The orbital dynamics of a satellite are highly nonlinear functions of its state. Therefore, approximation of the satellite state PDF as Gaussians could be quite a suboptimal *conjecture*. Knowledge about the orbit of a satellite is critical part of a space mission and has impacts on the power systems, attitude control and thermal design. Orbit Determination (OD) of a satellite in Low Earth Orbits (LEO) (orbit whose altitude from the surface of Earth ranges from 160 to 2000 km (100-1240 miles) [7]) is carried out using measurements from ground based sensors i.e., radars and onboard GPS device [8]. The measurements are also nonlinear function of the state of a satellite. In case of radar the measurements are only available once the satellite appears on the horizon, usually for 5-10 min. Moreover, these measurements are sometimes restricted due to an unsuitable satellite's orientation for strong return of radar energy. Contrarily, measurements from onboard Global

Positioning System (GPS) device are available throughout an orbit for LEO satellites. However, a satellite is equipped with limited power sources based on solar power and batteries [9]. Therefore, use of GPS device is required to be minimized in order to conserve power which directly influences space mission's life span. Thus, the measurements availability for OD of LEO satellites is mostly *sparse*.

In general sequential OD of a satellite for deep space endeavours such as mission to Moon also relies on fewer measurements. For example, consider a lunar orbiter optical navigation system. Its measurements could be angular quantities between stars and lunar surface landmarks. These measurements are nonlinear function of the state of a lunar orbiter. Moreover, their availability is only possible once the lunar surface landmarks and stars could be suitably viewed from the orbiter [1]. Therefore, full knowledge about time evolution or predictive PDF for satellite OD under *sparse* measurements becomes vital as it is used to quantify uncertainty associated with the state of a satellite until one receives the measurement. On receipt of measurement, the Bayes' formula is applied to update the *predicted* PDF based on *likelihood* of state. In practice to develop a practically realizable nonlinear filter there is a requirement of some tractable mathematical form for this PDF such as Gaussian approximation. Due to nonlinearity of dynamic and measurement systems in satellite OD problem, the use of Gaussian based nonlinear filters such as Extended Kalman Filter (EKF) [10],[5] is suboptimal. It is the most widely used nonlinear filter for sequential Bayes' filtering [5]. In EKF the system dynamics and measurement function are linearized to obtain suboptimal estimate and associated uncertainties. Due to linearization the region of stability could be small because nonlinearities in the system dynamics are not fully accounted. In plentiful measurement data environment, EKF could be considered sufficient for most real life requirements. However, there is a need for improvement in filtering techniques under *sparse measurement* data availability [11].

In addition to the state, a dynamical system may also depend upon parameters that are constant or perhaps known functions of time. The fundamental mathematical description of nonlinear satellite orbital dynamics is expressed in some Cartesian coordinate system (for details see Chapter. 3). The main forces affecting the orbit of a satellite are due to non-spherical Earth, atmospheric drag, gravitational attraction of Sun and other planets and radiation pressure [12]. In addition to the states of position and velocity of a satellite, the orbital motion also depends upon some parameters such as height of the orbit from surface of Earth and eccentricity of the orbit to name a few (for details see Chapter. 3). Apart from orbital parameters, the future form of an orbit in space is also characterized by some Initial Conditions (IC) provided to a satellite [13]. Given some suitable IC, the equations of motion are numerically integrated to obtain high precision satellite ephemerides. This is typically achieved by employing a very short time step to a numerical propagator. The calculation of the forces acting on a satellite at each time step slows down the computation which makes it prohibitive to use it on small satellites with less computational resources [14]. An alternative approach to numerical propagation of LEO satellites is use of analytic models [2]. Analytical orbit theories are very useful in

understanding and visualizing the perturbed description of an orbital motion [2],[15],[16]. For example recent interest in formation of spacecrafts in close proximity missions (separation distance of 250 - 500 m) like TanDEM-X [17] for Synthetic Aperture Radar (SAR) has revived the interest in understanding the description of relative motion of spacecrafts with each other and their long term perturbed orbital behaviour using analytical description of orbital motion. The theories could also help design orbit controller algorithm for constellation or formation maintenance and autonomous control [14]. However, in order to obtain an analytic solution the satellite's nonlinear equations of motions are linearized which makes the solution approximation of the true nonlinear dynamics. In general, the analytic solutions for an orbital motion are different from each other [2],[18],[19],[20],[15],[16]. This is due to dissimilar amount of approximation and linearization. Therefore, in order to use a particular analytical solution for actual space missions there is a requirement to analyze or investigate fidelity of that analytic model. Furthermore, in order to effectively utilize a particular analytic model proper selection of IC or parameters are crucial for their long term conformity to true nonlinear motion.

1.3 Discussion of Problem

The problem of Bayesian recursive filtering can be grouped into three types; (1) discrete, (2) continuous-discrete, and (3) continuous filtering [21]. The use of terms discrete and continuous denotes the way mathematical models of dynamic and measurement systems are expressed respectively. Filtering of a dynamical system where the system dynamics and measurement model are expressed in discrete time form is termed as discrete time filtering. These models are usually formulated as stochastic Discrete State Space Model (DSSM) owing to the way the system dynamics are propagated i.e., at fixed discrete instants and measurements also observed at discrete instants disturbed by additive white noise [5],[22]. The term stochastic appears due to uncertainties in physical effects and other disturbances modelled as white noise in DSSM. The evolution of time is a continuous process therefore dynamical systems can be more realistically represented as Stochastic Differential Equations (SDE) [6],[23]. In continuous-discrete filtering the term continuous represents progression of time continuously for system dynamics and discrete is used to represent measurements observed at fixed discrete instants [6]. Similarly to the DSSM, the continuous-time stochastic dynamic system is disturbed by an additive continuous time white noise and the measurements by a discrete-time white noise. The advantage of the continuous-discrete filtering is that the sampling interval can change between the measurements unlike discrete filtering where sampling time should be constant [21]. In continuous filtering, the system dynamics is represented as a SDE and the measurements are considered as a continuous-time process. An estimation problem is termed nonlinear if at least one model out of system dynamics or measurements is nonlinear. This work addresses nonlinear discrete and continuous-discrete type of filtering.

Probability theory provides a solution to recursive filtering problem as new observations are measured employing Bayes' formula [5]. Bayes' formula describes how PDF of the predicted state of a

dynamical system is changed based on the likelihood of current state of the system obtained from the measurement data. This is known as Bayes' *a posteriori* PDF. Considering the 1st order Markov property of the dynamical system, being addressed in this thesis disturbed by an additive white noise, the recursive form of Bayes' formula would require availability of *a posteriori* PDF of the state at a previous time only [23],[6],[5]. In the discrete-time filtering case this PDF is predicted forward using the total probability theorem known as Chapman-Kolmogorov-Equation (CKE) to obtain the predictive PDF [5]. A closed form solution for the CKE is only possible for linear systems for which the predictive PDF would be Gaussian [22]. In the continuous-discrete methodology the predictive PDF is obtained using the Fokker-Planck-Kolmogorov-Equation (FPKE) [24]. It is a linear Parabolic type Partial Differential Equation (PDE). The analytical solutions to this PDE are in general possible for linear dynamic systems only. Numerical solution for PDF of nonlinear dynamic systems is possible for low dimensions, ($\text{dimension} \leq 6$) due to recent increase in computational resources [25]. However, general use of numerical methods for solution of PDE in sequential filtering is not considered optimal [25] primarily due to their extensive computational aspects. The predictive PDF is updated using the *likelihood* of the current state using the Bayes' formula. Any optimal estimate criterion such as the Minimum Mean Square Error (MMSE) or *maximum a posteriori* (MAP) for the current state can be obtained from the Bayes' *a posteriori* PDF [22],[5]. Figure: 1-2 depict the block description of classic Bayesian recursive filtering methodology. Multidimensional integrals are employed to obtain MMSE or MAP estimates along with associated uncertainties in these estimates e.g., error covariance and higher order statistics from the Bayes' *a posteriori* PDF.

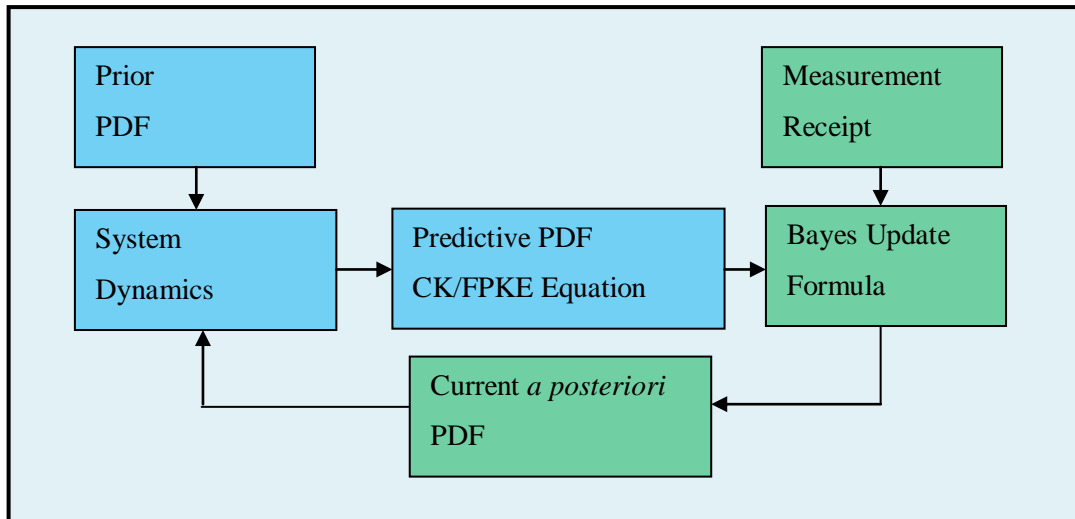


Figure 1-2: Block description of Bayesian prediction and update stages. The prior or *a posteriori* PDF of state at previous time is projected forward using CK or FPKE for discrete or continuous time dynamical system respectively. The predictive PDF is updated using Bayes' formula to obtain *a posteriori* PDF of state at current time.

In general for nonlinear dynamical systems such as satellite orbital dynamics the equations for mean and error covariance depends on all moments of Bayes' *a posteriori* PDF. However, this PDF cannot be characterized by finite set of parameters i.e., moments. Numerical solution of the Bayes' *a posteriori* PDF is in general intractable as it requires solution of CKE or FPKE which necessitates storage of the entire PDF. Therefore, one is forced to adopt approximations for the Bayes' *a posteriori* PDF. One would like to parameterize this PDF through a small set of parameters. If one is able to find a set of such parameters, a nonlinear filter would then comprise of equations for evolution of these parameters and consider these as sufficient statistics of the Bayes' *a posteriori* PDF. Nevertheless, it is practically impossible to find sufficient statistics for nonlinear problems [6].

There has been a considerable interest in approximating arbitrary non-Gaussian PDF using orthogonal expansions in terms of higher order moments of the distribution [26],[27],[28],[29]. Better approximations can be obtained by using more number of high ordered terms in such series expansions. An earlier approach of approximation for the Bayes' *a posteriori* PDF is orthogonal expansion of a Gaussian PDF in terms of higher order moments of the distribution and Hermite polynomials [1],[30]. Hermite polynomials are a set of orthogonal polynomials over the domain $(-\infty, \infty)$ with a Gaussian weighting function [31]. The resultant series is known as Gram Charlier Series (GCS) [29],[32],[28]. Previous work on use of such distributions for state estimation of nonlinear dynamical systems is restricted to single density expansion which has to be truncated at a particular low order moment term i.e., three in order to facilitate development of estimation algorithm [1],[33]. The use of GCS for Bayesian recursive filtering has shown improvement over EKF for nonlinear problems [1],[33]. However, the lower order expansions used in these references i.e., (order ≤ 3) are not optimal PDF approximations due to large deviation in centroid and negative probability regions [34]. Moreover, this type of PDF may not integrate to unity. There could be inference problems where single series may not be sufficient to model probability distributions especially multi-modalities [35]. Depending upon a particular type of PDF, higher order may be needed to obtain a good approximation in most of the cases. Increasing the order of series increases tremendous computational complexity and makes the series intractable especially for multivariate systems [28]. For example each increase in order adds $l = (o + d - 1)!/(o! (d - 1)!)$ moment terms where, o = order and d = multivariate dimension of PDF. Moreover, depending upon the type of the PDF to be approximated, the increase in such orders reach a certain point after which the approximation does not improve any further [36]. Recently, Van Hulle [34] suggested Gram Charlier Series Mixture Model (GCSMM) of moderate order expansion ($3 \leq \text{order} \leq 5$) to overcome difficulties associated with single series. Therefore, one may consider GCSMM of lower order GCS ($3 \leq \text{order} \leq 5$) as more optimal approximation of the Bayes' *a posteriori* PDF for state estimation of nonlinear dynamical systems.

Solutions of nonlinear differential equations obtained through numerical integration and their

analytical or linearized solutions are not exactly similar. In general this difference is time varying and termed as process noise [37],[38],[12],[5]. LEO satellite nonlinear models with forces due to non-spherical Earth gravitational potential, Atmospheric drag, luni-solar (Moon and Sun) gravitational attraction and solar radiation pressure increase complexity of equations of motion [12]. Numerical integration methods such as Runge-Kutta (RK) for solution of these equations can be employed to obtain high precision satellite trajectories for satellite state estimators and controllers [13]. However, numerical integration techniques are not suitable for On Board Computers (OBC) especially in small satellites due to resource limitations [14]. In general process noise for a particular analytical LEO model is exclusive. Propagation of orbital trajectories using analytical descriptions needs proper choice of orbital parameters or IC. The question arises how to choose IC of analytical approximation appropriate to a given choice for numerically propagated orbit obtained from nonlinear equations of motion such that the process noise is minimized. This would entail two trajectories to be sufficiently close to each other. Furthermore, it provides an insight into fidelity of an analytical model and their long term perturbed orbital behaviour.

1.4 Aims and Objectives

1.4.1 Aims

In view of the nonlinear estimation problem the aims of this research are as under:

1. Develop sequential Bayesian filters for nonlinear dynamical systems.
2. Analyse and compare fidelities of linearized LEO orbital models.
3. Estimate parameters for analytic orbital model [2] around the oblate Earth.

1.4.2 Objectives

The above aims are translated into following objectives:

1. Develop sequential Bayesian filters for nonlinear dynamical systems in general and satellites in particular using GCS and GCSMM and simulate their performance under *sparse measurements* availability.
2. Analyse and investigate process noise of linearized LEO absolute and relative motion orbital models, with a view to compare their fidelities, using Gauss-Legendre-Differential-Correction (GLDC) method.
3. Develop high precision Epicyclic orbit [2] parameter filter based on linear least squares [38].

1.5 Structure of Thesis

The research presented in this thesis is focused on both parameter and state estimation of nonlinear

dynamical systems in general and LEO orbital dynamics in particular. It consists of seven chapters. Chapter: 2 present literature survey on parameter and state estimation of dynamical systems and LEO orbital mechanics. Chapter: 3 elaborates on analysis of fidelities of linearized orbital models for LEO using GLDC method [39][40]. Firstly, two absolute orbital motion models i.e., Epicycle Model for Oblate Earth [2] and Kepler's 2 *body problem* [13] are analyzed. Secondly, analysis of two analytical models describing relative motion of spacecrafts with each other i.e., Hill-Clohessy-Wiltshire (HCW) equations [18],[19] and Schweighart and Sedwick (SS) J_2 modified Hill's equations [20] is carried out. Chapter: 4 presents the Epicycle orbit parameter filter using linear least squares [38]. Initially a brief description of the Epicycle model is presented which focuses on key idea used in the filtering algorithm. The algorithm exploits linear secular terms in Epicycle coordinates of argument of latitude and right ascension of the ascending node. Accurate determinations of orbital parameters enable high fidelity long term orbital propagations. Chapter: 5 present GCS and its Mixture Particle Filtering. Firstly, it investigates generic Particle Filters (PF) [41], Gaussian Particle Filters (GPF) [42] and Gaussian Sum Particle Filters (GSPF) [43]. Subsequently, it develops a PF based on GCS and its Mixtures. The filtering algorithms are simulated on nonlinear simple pendulum model and OD of spacecraft in LEO orbits. Chapter: 6 present the Kalman [10],[6] and Culver Filter (CF) [1] frameworks for Bayesian filtering of nonlinear dynamical systems. The Kalman Filter framework consists of the EKF and Gaussian Sum Filter (GSF) [44]. The Culver framework constitutes of third order CF and its new extension called Mixture Culver Filter (MCF) [35]. Firstly, the algorithms used in Culver frameworks are described in detail. Subsequently, the algorithms are simulated and analyzed for radar and GPS based OD of a satellite in LEO orbits and optical navigation for a lunar orbiter [1]. Chapter: 7 present future research directions and conclusion.

1.6 Novelty

The contributions of this thesis are summarized below:

- Based on MC simulation approach [41],[45],[42], new GCS / GCSMM particle filters and hybrids are developed for nonlinear Bayesian discrete-time state estimation. The use of such PDFs for nonlinear estimation under *sparse measurements* availability has shown improvement over other filtering methods such as EKF and generic Particle Filter (PF).
- Based on Taylor series expansion of nonlinear dynamic equation and third order GCSMM approximation of the Bayes' *a posteriori* PDF a new nonlinear filter namely MCF is developed. This approach is essentially an extension of an earlier work by Culver [1] (in this thesis it is termed as Culver Filter (CF)). MCF serves as an exact solution to Bayesian filtering problem. More notably it utilizes optimal FPKE error feedback to compute certain parameters of GCSMM associated with each of its component.
- The application of new nonlinear Bayesian filters based on GCS and GCSMM are

simulated for simple pendulum, LEO satellite OD and navigation of lunar orbiter under *sparse measurements* and compared with other state of the art nonlinear filters such as EKF. This provides a unified investigation on use of GCS and GCSMM for nonlinear state estimation based on Taylor series and MC simulations.

- A new analysis on fidelities of linearized LEO absolute and relative motion orbital models using GLDC scheme [46],[39],[40]. The selection of appropriate IC or parameters of analytic models is imperative to minimize the process noise and obtain more accurate orbital trajectories.
- A new algorithm based on linear least squares for parameter estimation of Epicyclic orbit is developed. The estimator is termed as Epicycle Parameter Filter (EPF). The method exploits the linear secular increase in Epicyclic coordinates. The estimated parameters enable minimization of the process noise and long term high fidelity orbital trajectory generation at all inclinations for LEO [38].

1.7 Publications

List of publications is as under:

- “Analysis of Fidelities of Linearized Orbital Models using Least Squares” by Syed A A Gilani and P L Palmer presented at IEEE Aerospace Conference 2011, 5-12 Mar 2011 at Big Sky, Montana, USA.
- “Epicycle Orbit Parameter Filter for Long Term Orbital Parameter Estimation” by P L Palmer and Syed A A Gilani presented at 25th Annual AIAA/USU Conference on Small Satellite 8-11 Aug 2011 at Logan, Utah USA.
- “Nonlinear Bayesian Estimation Based on Mixture of Gram Charlier Series” by S A A Gilani and P L Palmer, presented at IEEE Aerospace Conference 2012, Mar 2012 at Big Sky, Montana, USA.
- “Sequential Monte Carlo Bayesian Estimation using Gram Charlier Series and its Mixture Models”, by S A A Gilani and P L Palmer, proposed for IEEE Journal of Aerospace (write up is in progress)

2 Literature Survey

2.1 Nonlinear Bayesian Recursive Filtering

Nonlinear filtering has been a subject of an immense interest in the statistical and other scientific community for more than fifty years [6],[1]. The central idea of Bayesian recursive filtering is availability of Bayes' *a posteriori* PDF based on all available information about the dynamical and measurement systems and prior knowledge about the system [5],[47]. One may satisfy the optimality criterion of the MMSE or MAP for current state estimates and their error statistics from this PDF. In general, a tractable form of the Bayes' *a posteriori* PDF is difficult to obtain except for a limited class of linear dynamical and measurement systems. In practice approximate forms of this PDF are used instead. These methods can be broadly grouped into: (1) Gaussian based methods, [10],[48],[42] (2) Gaussian Mixture Model (GMM) based methods, [44],[49] (2) Sequential Monte Carlo (SMC) methods, [41],[45],[50],[47] (3) Orthogonal Expansion based methods, [33][30] (4) Numerical methods, [8],[51] and (5) Variational Bayesian methods [52]. In the subsequent sections a review of each of these approaches will be presented.

2.1.1 Gaussian Based Methods

In order to obtain the Bayes' *a posteriori* PDF and compute MMSE or MAP estimates one would require moments of the *a posteriori* PDF. These are integrals over an infinite domain $(-\infty, \infty)$ [5],[6]. It is usually difficult to obtain tractable forms of the PDF required for analytical expression of integrals. Moreover, such solutions, if obtained through numerical integration would require storage of the entire PDF which is an infinite dimensional vector [5]. In linear systems the Bayes' *a posteriori* PDF is considered to be Gaussian for which the Kalman Filter (KF) is the optimal MMSE or MAP solution [10]. The use of KF equations for nonlinear filtering is made possible by linearizing the dynamic and measurement equations to obtain an approximate filtering method, known as EKF [6]. In the EKF one computes only the first two moments i.e., mean and variance of Bayes' *a posteriori* PDF. Therefore, it is commonly termed as a Gaussian method for filtering of nonlinear systems [22]. In such applications it could produce very erroneous estimates, for example it computes expected value of a function $\mathbf{f}(\mathbf{x})$ as $E[\mathbf{f}(\mathbf{x})] \approx \mathbf{f}(E[\mathbf{x}])$ which is true only for linear functions. For example, consider a nonlinear function $\mathbf{f}(\mathbf{x}) = \mathbf{x}^2$. If one considers the mean of \mathbf{x} to be zero, this would give the following EKF approximation $E[\mathbf{x}^2] \approx (E[\mathbf{x}])^2 = 0$, whereas the true value of the variance $E[\mathbf{x}^2]$ could be any positive value [25]. However, an important historical significance of the EKF is its use for Guidance and Navigation for the Apollo mission to the Moon [53]. Recently new nonlinear

filtering methods based on deterministic sampling of the Bayes' *a posteriori* PDF have emerged to improve the performance of the EKF. The first such algorithm was introduced by Julier and Uhlmann known as Unscented Kalman Filter (UKF) [48]. There have been many improvements of the UKF. The class of such filters is collectively known as Sigma Point Kalman Filters (SPKF) [22]. The SPKF uses a set of deterministically weighted sampling points known as "sigma points" to parameterize the mean and covariance of a probability distribution for a nonlinear system considered as Gaussian. The sigma points are propagated through nonlinear systems without any linearization unlike the EKF. These filters avoid the explicit computation of Jacobian and/or Hessian matrices for nonlinear dynamic and measurement functions. Therefore, these filters are commonly termed as derivative free filters. Derivative free filters have a distinct advantage through their ability to tackle discontinuous nonlinear dynamic and measurement functions. Two important closely related algorithms are the Central Difference Filter (CDF) [54] and Divided Difference Filter (DDF) [55]. These filters employ an alternative linearization approach for the nonlinear functions. The approach is based on the Stirling's interpolation formula [56]. Similar to the UKF these algorithms are based on a deterministic sampling approach and replace derivatives with functional evaluations. Merwe [57] improved these algorithms to provide computationally more reliable square root versions known as Square-Root UKF (SR-UKF) and Square-Root CDF (SR-CDF) [58]. Use of SPKF for satellite orbit determination is considered in [59].

2.1.2 Gaussian Mixture Model Based Methods

Any non-Gaussian PDF can be approximated as a linear sum of Gaussian PDFs known as GMM [60]. Complex PDF structures such as multiple modes and highly skewed tails can be efficiently modelled using a finite GMM. In the seminal work of Alspach, the GMM is used to approximate Bayes' *a posteriori* PDF in nonlinear filtering applications [44]. This nonlinear filter is called Gaussian Sum Filter (GSF). It is essentially a bank of parallel running EKF to solve the Bayes' sequential estimation problem. The mean and covariance of each individual Gaussian component is updated using the EKF methodology. Therefore, the GSF could also suffer from reduction in region of stability due to the use of the EKF as a basic building block. However, it has shown improvement over the EKF in nonlinear filtering applications [44],[61]. Furthermore, the concept of GMM for the Bayes' *a posteriori* PDF has been used to develop the Gaussian Mixture Sigma Point Particle Filter (GMSPPF) [22] and Gaussian Sum Particle Filtering (GSPF) [43]. In the GMSPPF the use of an EKF has been replaced with sampling based filters i.e., UKF or CDKF to obtain the mean and covariance of each Gaussian component; whereas, in the GSPF Monte Carlo (MC) simulation [41] is used to obtain these parameters. A further improvement of the GSF is reported in [49] where weight updates for GMM are obtained using the error feedback acquired based on minimizing the Integrated Square Error (ISE) for the predictive filtering PDF solved by the FPKE and a filter generated GMM approximation. Nonlinear filters based on GMM are computationally more expensive. Keeping the number of GMM

components fixed in nonlinear filters could be a suboptimal representation for a continuously evolving Bayes' *a posteriori* PDF. To overcome this problem an adaptive GMM has been suggested in references [62],[63].

2.1.3 Sequential Monte Carlo Methods

Another recent approach to find solutions to the Bayesian inference problem is through MC simulations [47]. A recursive form of the MC simulation based on a Bayesian filtering scheme is known as Sequential Monte Carlo (SMC) method. In SMC method restrictive assumption of linear DSSM and Gaussian Bayes' *a posteriori* PDF is relaxed. A set of discrete weighted samples or particles are employed as point mass approximations of this PDF [41],[22],[64]. The point masses are recursively updated using a procedure known as Sequential Importance Sampling and Resampling (SIS-R) [41]. The SIS-R is a process in which particles are sequentially drawn from a known easy to sample proposed PDF considered as approximation to the true Bayes' *a posteriori* PDF. The point mass approximation of PDF in this filter leads a summation form of Bayesian integrals. Therefore, MMSE or MAP state estimates and associated uncertainties are conveniently obtained. Due to their ease of implementation and ability to tackle nonlinear DSSM, its use is found in various diverse applications [59],[65]. This nonlinear Bayesian filter is termed as Bootstrap or Particle Filter [41].

The generic Particle Filter (SIS-R) has undergone a number of improvements since its development. A serious shortfall affecting particle filters is their lack of diversity or degeneration of particles. This is because the proposed PDF does not effectively represent the true Bayes' *a posteriori* PDF. Therefore, one may consider an EKF or a SPKF to generate a better approximation of the Bayes' *a posteriori* PDF which can be used for the proposal PDF [22],[50],[45]. Generic particle filters do not assume any functional form for the predictive or Bayes' *a posteriori* PDF. However, a consideration could be Gaussian or GMM forms for these PDFs [42],[43]. Accordingly, sampling of particles is carried out using the assumed PDF. In this thesis an extension to these methods are developed employing GCS and its mixture models. Sequentially sampling and resampling from a discrete proposed PDF in SIS-R produces sample degeneration and impoverishment. In order to overcome this problem a continuous time representation of the Bayes' *a posteriori* PDF is introduced in the particle filter known as Regularized Particle Filter (RPF) [66]. Kernel PDF estimation methods [67] are employed to obtain a continuous time representation of Bayes' *a posteriori* PDF. Typically, Epanechnikov or Gaussian Kernels are employed for such estimation methods [68]. Resampling from approximate Bayes' *a posteriori* PDF is carried out using the continuous time representation. A closely related filter named as the Quasi-Monte Carlo method implements Bayes rule exactly using smooth densities from exponential family [69].

In multivariate nonlinear filtering, estimation problems can occur in which one may partition the state vector to be estimated, depending upon a particular DSSM. The partitioning is based on components

of the state space which can be estimated using analytical filtering solutions such as Kalman Filter [10] and the components which require nonlinear filtering methods such as SIS-R [70],[71]. The fundamental idea is to develop recursive relations for a filter by decomposing Bayes' *a posteriori* PDF into one generated by a Kalman Filter and the other formed by a SIS-R particle filter. This hybrid filtering method is known as Rao-Blackwell Particle Filter (RBPF). The RBPF for higher dimensional state vectors with fewer particles is expected to give better results compared with high number of particles for a SIS-R [8].

In general high fidelity measurement systems have low noise levels compared with the dynamic system noise. Therefore, Bayes' *a posteriori* PDF is likely to resemble more with the likelihood compared with the proposed PDF used in SIS-R. Particle filtering of such systems can be improved by considering the likelihood function as the proposed PDF [68]. Pitt and Shephard introduced a variant of a SIS-R particle filter by introducing an auxiliary variable defining some characteristic of the proposed PDF e.g., the mean [72]. This filter is known as Auxiliary sampling importance resampling particle filter. The difference between a generic SIS-R and this filter is at the measurement update stage where the weights of each particle would be evaluated in the latter using parametric conditioning of the likelihood [68].

2.1.4 Orthogonal Expansion Based Methods

There has been a considerable interest over a long period of time in the use of orthogonal expansions of the PDF for analysis and modelling of non-Gaussian distributions, among statistics community [32],[29],[73],[74]. Use of Hermite polynomials for expansion of Gaussian PDFs in terms of higher order moments of a particular distribution is well known as GCS or Edgeworth Series [28],[29]. Hermite polynomials are a set of orthogonal polynomials over the domain $(-\infty, \infty)$ with Gaussian weighting function $(e^{-x^2/2})$ [27],[31]. The ability of GCS to model non-Gaussian distributions has led researchers in nonlinear estimation and Bayesian statistics to develop nonlinear filtering algorithms based on GCS approximation of Bayes' *a posteriori* PDF [1],[75],[33],[76],[30]. In 1969 Culver developed closed form analytical solutions for the nonlinear Bayesian inference problem using third order GCS to approximate predictive and Bayes' *a posteriori* PDF for a continuous-discrete nonlinear filtering scheme [1]. In this nonlinear filter, instead of using FPKE to obtain predictive PDF, higher order moments of the distribution are used to formulate its GCS approximation. However, the linearization of dynamic and measurement models is carried out to facilitate the filter development. In this thesis this filter will be named as Culver Filter (CF). Apart from the analytical solution of integrals involving exponential series, the use of GCS is convenient for numerical integration technique such as Gauss Hermite Quadrature (GHQ) [77]. In GHQ the numerical computation of such integrals is considerably reduced as evaluation of integrands is only done at deterministically chosen weighted points. These points are roots of the Hermite polynomials used in GCS. In nonlinear

filtering, the GHQ method for solution of Bayesian inferences has also been extensively employed [33],[30],[76]. Challa [33] developed a variant of CF using a higher order moment expansion of the predictive PDF, very similar to the one developed by Culver. However, in that filter the Bayes' formula was solved numerically using GHQ with weighted points obtained from an EKF (or Iterated EKF [5]). In general, GHQ can also be used for computing coefficients of the GCS also known as Quasi-Moments [1] and develop approximation for Bayes' *a posteriori* PDF [30],[76]. Horwood developed an Edgeworth filter for space surveillance and tracking using a GHQ based numerical solution of Bayesian integrals [62]. In this thesis a GCS based nonlinear filters have been developed using SMC scheme [47]. Moreover, extensions based on GCSMM are developed for nonlinear discrete time and continuous-discrete filtering.

2.1.5 Numerical Based Methods

The Nonlinear filtering methods discussed so far in this chapter approximate Bayes' *a posteriori* PDF with Gaussian, GCS or point mass PDF approximations. However, numerical methods for the solution of differential and integral equations can be used to obtain close to exact Bayes' *a posteriori* PDF and associated inferences [5]. Conceptually, in nonlinear filtering one has to solve the FPKE or CK (discrete filtering case) to obtain the predictive PDF. The Use of numerical methods for solution of FPKE especially for the multi-dimensional case is prohibitive due to excessive computations. The solution of such a PDE is described on a fixed grid in a d -dimensional space (where, d = number of dimensions). The computational complexity increases as N^d (where, N = number of grid points in each dimension) [25]. Kastella and Lee developed nonlinear filters based on Finite Difference (FD) method [78] for numerical solutions of 4-dimensional FPKE [8],[51]. A closely related method exists for discrete time filtering known as Grid Based Filters (GBF) [68]. The GBF approximates Bayesian integrals with large finite sums over a uniform d -dimensional grid that encompasses the complete state space of a nonlinear dynamic system. Another relatively new concept of approximating the PDE is a mesh free method which utilizes an adaptive grid instead of a fixed grid [79]. Mesh-free methods are considered as better solutions for the FPKE equation compared with the SIS-R particle filter generated point mass PDF approximation. It is due to the inherent smoothness of PDE solutions [25]. An integrated nonlinear filter based on offline numerical solution of FPKE and Kalman filter has been developed by Daum [80]. The filter could also handle diffusions belonging to the exponential family like the Maxwell-Boltzmann distribution contrary to usual Gaussian type diffusions [6].

2.1.6 Variational Bayesian Methods

Variational Bayesian (VB) methods are commonly known as “ensemble learning”. These comprise a family of new methods to approximate intractable Bayesian integrals thereby serving as an alternative to other approaches discussed above. In these methods the true Bayes' *a posteriori* PDF is approximated by a tractable form, establishing a lower or upper bound. The integration then forms

into a simpler problem of bound optimisation making the bound as tight as close to the true value [52]. A lower bound of the likelihood of a *a posteriori* PDF is maximized with respect to parameters of the tractable form using Jensen's inequality and variational calculus.

2.2 Parameter Estimation

In addition to the state, a dynamical system may also depend upon parameters that are constant or perhaps known functions of time, for example the mass of bodies in a mechanical model or the birth rate and carrying capacity in a population model. In addition to the state of angular position and velocity of a simple pendulum the model also depends upon two parameters, the pendulum's length and the strength of gravity. The parameter is typically a time-invariant vector or a scalar quantity of a particular dynamical system. A parameter could govern a qualitative behaviour of the system, such as a loss of stability of its solution or a new solution with different properties. One may also consider it to be slightly time varying but its time variation is slow compared with the state estimation discussed earlier in this chapter. Parameter estimation could be performed with two main approaches, Bayesian or Non-Bayesian [5].

In Bayesian approach, one seeks Bayes' *a posteriori* PDF of parameters using Bayes' formula. The MMSE estimates are obtained as mean, and MAP as mode, of Bayes' *a posteriori* PDF. In the non-Bayesian approach no prior assumption on the type of probability distribution of the parameters is made. However, one may utilize a *likelihood* function which is the probability distribution of the measurements conditioned on the parameter of interest. The estimate obtained by maximizing the *likelihood* function with respect to the parameter of interest is the *Maximum Likelihood Estimate* (MLE) [5].

In least squares method sum of the squares of the errors between the measurement obtained from measurement system and the modelled dynamics are minimized with respect to the parameter of interest [5],[81]. There is no assumption on probability distribution of these errors. Recursive and non-recursive least squares (without process noise) were both invented by Gauss. Due to the nonlinearity of celestial mechanics laws, he used linear approximation for the dynamics just like in the EKF [25]. If the measurement errors are assumed independent and identically distributed (i.i.d) with the same marginal PDF, zero-mean Gaussian distributed, then the method coincides with the MLE. There is a large literature devoted to these methods in almost all fields of physical sciences and engineering including astrodynamics [81],[12] tracking and navigation [5]. In this thesis nonlinear least squares commonly known as GLDC [12],[82] is considered for the analysis of fidelities of linearized LEO orbital models [37].

2.3 Satellite Orbital Dynamics

The orbital motion of a satellite around the Earth is described in its simplest form found out empirically by Kepler about 400 years ago [83],[84]. The acceleration of the satellite in a

gravitational field is given by the Newton's law of gravity [13]. The problem of satellite motion under the influence of spherically symmetric Gravitational potential is known as the *2 body problem* [14]. It has been shown that the satellites orbit in conic sections depends upon the total energy of the satellite [12],[14]. Such an orbit has an angular momentum and orbital energy both conserved and is called a *Keplerian orbit* [85]. The satellite remains in a fixed plane around the Earth, called the orbital plane. An orbital plane is defined by the position and velocity vectors of a satellite. In reality forces due to the Earth's non-spherical shape, atmospheric drag (especially for LEO), Gravitational attraction of Sun and Moon, and solar radiation pressure significantly influence the orbit of a satellite. The effect of the non-spherical Earth on the orbits of a satellite has been studied extensively for about 50 years. The trajectories of satellite orbits are expressed in terms of instantaneous Keplerian orbital elements or osculating elements. Gauss's planetary equations of motion describe the evolution of these orbital elements under perturbing forces [12]. Kozai [15] and Brower [16] found analytic solutions to perturbed orbital elements of a satellite. Hashida and Palmer [2] developed a simplified and accurate analytical description of the satellite orbital motion around an oblate planet [2]. The model has a simple analytic form and is capable of describing all the gravitational perturbative effects. The formulation of the orbit is based on the Epicyclic motion of a satellite [14]. In order to find out an analytical solution for equations of motion of satellite dynamics with additional forces, one needs to linearize about some reference satellite trajectory usually a circular or low eccentricity orbit. The analytic propagation of orbits is not very accurate for long duration of time when compared with the true satellite trajectories or even from the numerical integration of the complete equations of motion [15],[16],[14].

2.4 Satellite Relative Motion

Satellite formations have received extensive attention for global observation [17], communication [86] and stellar interferometry [87], due to advantages of flexibility and low cost. The description of motion of a satellite flying in a formation is determined from the relative motion of two satellites. In this scenario, one of the satellites known as the deputy satellite is considered in a relative coordinate system fixed to another satellite known as the chief satellite [88]. The deputy satellite's relative motion can be conveniently described by subtracting the absolute motion of two satellites (chief and deputy) around the planet. HCW equations provide the simplest model describing satellite relative motion [18],[19]. These are second order linearized differential equations describing relative motion of a satellite in a near circular orbit around a spherical Earth. Extensive work on improvement of HCW equations to include the effects of oblate Earth, eccentric orbit and nonlinear differential gravity acceleration has been carried out. The linearized differential equations describing the relative motion in unperturbed elliptical chief orbit were presented by Lawden [89] and Tshauer-Hempel [90]. In reality, due to non-spherical geopotential an orbit of chief satellite would experience secular and periodic changes in its orbital elements over time. Secular variations in a particular element change

linearly over time and cause unbounded error growth. Periodic changes are of two types: (1) short periodic, and (2) long periodic, depending upon the amount of time required for the effect to repeat. Short periodic effects repeat on the order of satellite orbital period or even less. Long periodic effects have cycles significantly larger than one orbital period which are usually one or two times longer than short periodic [12],[2]. Therefore, satellite relative motion models based on spherically symmetric geopotential for chief orbit such as HCW equations are not good approximations over the whole period of time. Schweighart and Sedwick (SS) [20] derived equations for the relative motion between satellites in a formation, incorporating the effects of Geo-potential zonal harmonic J_2 . The relative motion of satellites based on Epicyclic model was presented by Halsall and Palmer [91]. Due to approximations of chief satellite motion, all of the analytical or approximate models described above have varying levels of process noise. Therefore, there is a need to compare the validity and usefulness of these models over time. This requires some methodology to find the most suitable approximate orbit to use in this comparison.

2.5 Summary

This chapter presents a brief literature review on the main approaches for parameter and state estimation of dynamical systems. LEO absolute and relative orbital mechanics have also been briefly discussed with a view to develop understanding of the major difference between true dynamical model and analytical model. In particular, the estimators have different performance for different dynamical systems. The accuracy and strength of the simplifying assumptions in different algorithms strongly depends upon the inference problem at hand which makes a particular approximate solution better than others. In nonlinear filtering applications where less measurements are available, such as radar based orbit determination of a space object, would necessitate good prediction accuracies. In this application radar measurements of a space object are available for a very short span of time i.e., during the time satellite is visible on the horizon, typically 5-10 min for a LEO satellite. It requires better understanding of the nonlinear dynamical system and an approximation for its time varying probability distribution i.e., non-Gaussian. Therefore, Gaussian based assumptions for state predictive and Bayes' *a posteriori* PDF for nonlinear dynamical systems like satellite orbital dynamics would be suboptimal. Thus, GCS and GCSMM are employed in this thesis for state predictive and Bayes' *a posteriori* PDF for filtering of nonlinear dynamical systems under less measurements accessibility.

3 Analysis of Fidelities of Linearized Orbital Models

3.1 Introduction

The dynamics of a satellite orbiting the Earth is described by second order nonlinear differential equations. These nonlinear equations do not have an analytical solution except for the *2 body* problem of a satellite around a spherically symmetric gravitational potential [13]. However, a satellite around our Earth is subjected to additional forces due to *non-spherical* Earth, atmospheric drag, gravitational attraction of other heavenly bodies like the Moon and Sun (Luni-Solar) and solar radiation pressure [14]. These forces are termed as perturbations to a *Keplerian orbit* as they are all much smaller than the acceleration due to spherical Earth [83],[84],[14]. Table: 3-1 lists important sources of perturbations, and their effects in terms of accelerations acting on satellite as a function of its orbit height above Earth [14]. The perturbation due to geopotential second zonal harmonic term known as J_2 (explained later in this chapter) is the most dominant perturbative effect on a satellite, which is due to oblateness of Earth [12].

Source	$h = 150 \text{ km}$	$h = 750 \text{ km}$	$h = 1500 \text{ km}$	$h = 35785 \text{ km}^*$
Spherical Gravity	9.35	7.85	6.42	0.22
Earth Oblateness J_2	30×10^{-3}	20×10^{-3}	14×10^{-3}	160×10^{-7}
Atmospheric Drag	3×10^{-3}	10^{-7}	-	-
Luni-Solar Attraction	10^{-6}	10^{-6}	10^{-6}	70×10^{-7}
Solar Radiation Pressure	10^{-7}	10^{-7}	10^{-7}	10^{-7}

h = height of orbit above Earth

*Satellite at this height is called *geosynchronous* because its orbital period around the Earth matches the rotation rate of Earth around its polar axis which is 23 hours 56 min and 4 sec [92].

Table 3-1: Disturbing Forces on Satellites in m/s^2

Extensive analysis has been carried out since the dawn of satellite age to study the motion of a satellite under a *non-spherical* gravitational potential [2]. These equations of motion are highly nonlinear and analytic solutions are only available for linearized forms. Such solutions are

advantageous in visualizing the long term perturbed orbital behaviour. They provide useful insight into the physics of the orbital motion. However, due to linearization of equations of motion these solutions are considered as approximations of the full nonlinear dynamical model. The difference between the nonlinear equations of motion and their approximation is termed as process noise. In general, the process noise in an analytically derived model is a time varying quantity. It depends directly upon the approximations and assumptions applied on the nonlinear dynamic model. This varies the fidelities of analytical models and impacts upon their use for modelling actual space missions. It is usually very difficult to model process noise with some fixed parameters. For example in nonlinear filtering such as EKF if one uses a particular analytical orbital model [2],[15],[16],[93], the un-modelled accelerations and effects due to linearization are modelled as white noise [5]. In practice, the parameters of such a noise are approximated as moments (up to second order) of a Gaussian PDF [5]. These parameters would require adjustments for optimum performance of filters. In general, an analytical model with superior fidelity would yield better results with such adjustments. Moreover, in satellite orbit control applications the use of analytical models can assist the establishment of orbit controller algorithms [94],[95],[96],[97],[98]. Therefore there is a need to carry out qualitative and quantitative analysis of linearized orbital models and their process noise. Linearized solutions are characterized by a set of IC which determines the orbital evolution. If their IC is not properly chosen, the evolution would sooner or later diverge as shown in Figure: 3-1. This results in their validity for a very short period of time.

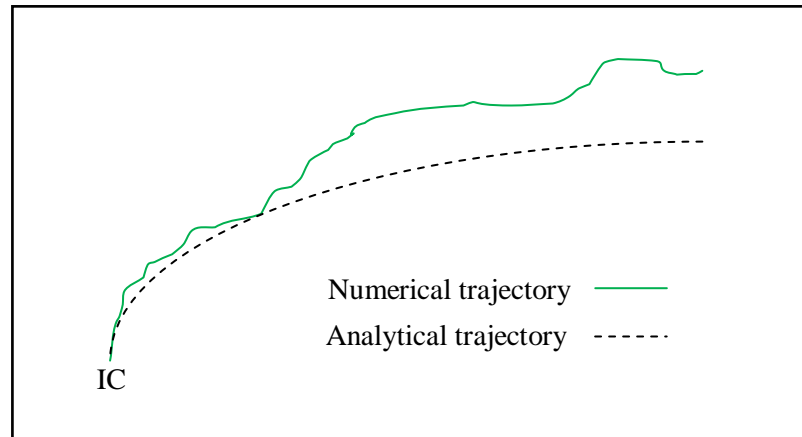


Figure 3-1: The concept of divergence. Numerical integration of nonlinear equations of motion is termed as “Numerical trajectory” and the linearized solution as “Analytical trajectory”. This picture shows limited time validity of the analytical solutions when compared with a numerically obtained trajectory due to the process noise. The choice of IC is very critical as it determines the form of the motion at later times.

The question arises on how to choose the IC of the analytical approximation appropriate by a given choice of IC for the numerically integrated nonlinear equations of motion such that the process noise is minimized. As process noise is a peculiar quantity for a particular analytical model. Therefore, the amount of its minimization is also unique. In general, the minimization of the process noise enables us to obtain the approximate trajectory close to the numerical nonlinear trajectory and provides useful insight into fidelity of an analytical model and assists in comparing validity and usefulness of these models over time.

Keeping in view the discussion so far, analyzing fidelities of analytical models and minimizing their process noise is certainly an interesting and valuable research. In this chapter the GLDC scheme [46],[12] is adapted as a solution methodology for this analysis. Firstly, the solution methodology for our analysis is developed followed by a brief description on fundamentals of orbital mechanics. Secondly, the analysis of LEO satellite absolute and relative orbital models will be presented.

3.2 Methodology for Fitting Approximate Models to Nonlinear Data

The GLDC is a useful statistical method for satellite orbit determination which dates back to Gauss (1801) [39],[46]. The quantities of interest in satellite orbit determination could be the orbital parameters or position and velocity of a satellite in some Earth based coordinate system. These quantities are usually measured indirectly using sensors. In a GPS based orbit determination, the position and velocity of a satellite are measured in a co-rotating coordinate system fixed to the Earth, using an onboard GPS sensor; whereas, in radar based orbit determination the measured quantities are: (1) radar site to satellite position vector (ρ), angular quantities of (2) azimuth (Az), and (3) elevation (El) of radar antenna in a radar site based coordinate system (details in Chapter 5) [8]. Both types of measurements are nonlinear functions of the position and velocity of a satellite in a coordinate system fixed to some reference direction in space [12]. Unlike the orbit determination problem, consider that position and velocity of a satellite are directly available by numerically integrating nonlinear equations of motion, using some numerical method such as Runge-Kutta (RK-4) [78], without using sensors. This means that the satellite orbital trajectory in a particular coordinate system is fully known without any measurement noise. Let each of this position and velocity vector in a three dimensional coordinate system of Earth are accessible for a specific period of time are expressed as:

$$\mathbf{Z}(\mathbf{x}_0, t) = [(\vec{r}_0, \vec{v}_0), (\vec{r}_1, \vec{v}_1), \dots, (\vec{r}_k, \vec{v}_k)]^T \quad (3.1)$$

where, $\mathbf{x}_0 = (\vec{r}_0, \vec{v}_0)^T$ is the vector of initial conditions for $3 \times$ position (\vec{r}_0) and $3 \times$ velocity (\vec{v}_0) coordinates and k = time subscript of the numerically computed position and velocity vectors of a satellite with three components each. This forms as a large vector of nonlinear data of $6k$ components for the differential correction scheme. The term “reference or numerical trajectory” will be used for

the nonlinear data. The position and velocity vectors of a satellite approximate analytic model for each of the corresponding time instant as in Equation: 3.1 can be obtained by providing the IC to the analytic model and expressed as:

$$\mathcal{P}(\mathbf{x}_0, t) = [(\vec{r}_0, \vec{v}_0), (\vec{r}_1^p, \vec{v}_1^p), \dots, (\vec{r}_k^p, \vec{v}_k^p)]^T \quad (3.2)$$

where, p is the superscript denoting the position and velocity vectors of an approximate analytic model. The term “analytical trajectory” will be used for the approximate analytic model data. Note the IC \mathbf{x}_0 provided to approximate analytic model is the same as in case of full nonlinear equations of motion. However, the trajectory at later time would be different or diverged due to the process noise. Consider that the nonlinear equations of motion have been accurately modelled by the analytical approximation less process noise errors $\mathbf{q}(t)$. Therefore, one may write:

$$\mathbf{q}(t) = \mathbf{Z}(\mathbf{x}_0, t) - \mathcal{P}(\mathbf{x}_0, t) \quad (3.3)$$

where, $\mathbf{q}(t)$ is assumed to be as independent distributed Gaussian random variable with mean $\tilde{\mu} = 0$ and variance σ^2 . Therefore, the expression for the variance may be written as:

$$\sigma^2 = E[(\mathbf{q}(t))^2] \quad (3.4)$$

where, $E[.]$ is the expectation operator [99]. Using Equation: 3.3, the variance can be rewritten as:

$$\sigma^2 = \frac{1}{6k} (\mathbf{Z} - \mathcal{P})^T (\mathbf{Z} - \mathcal{P}) \quad (3.5)$$

where, $\mathbf{Z} = \mathbf{Z}(\mathbf{x}_0, t)$ and $\mathcal{P} = \mathcal{P}(\mathbf{x}_0, t)$. In order to minimize the difference between the orbit $\mathbf{Z}(\mathbf{x}_0, t)$ and $\mathcal{P}(\mathbf{x}_0, t)$, we consider the variance σ^2 as a cost function to be minimized. The procedure is to differentiate the cost function with respect to \mathbf{x}_0 and set it equal to zero:

$$\frac{\partial \sigma^2}{\partial \mathbf{x}_0} = \left(\frac{\partial \mathcal{P}}{\partial \mathbf{x}_0} \right)^T (\mathbf{Z} - \mathcal{P}) = 0 \quad (3.6)$$

where, following vector derivative relation is used in deriving Equation: 3.6 [13]:

$$\frac{\partial \mathbf{a}^T \mathbf{b}}{\partial \mathbf{c}} = \mathbf{a}^T \frac{\partial \mathbf{b}}{\partial \mathbf{c}} + \mathbf{b}^T \frac{\partial \mathbf{a}}{\partial \mathbf{c}} \quad (3.7)$$

where, \mathbf{a} , \mathbf{b} and \mathbf{c} are vectors. Based on the assumption of differentiability of equations of motion the analytic trajectory $\mathcal{P}(\mathbf{x}_0, t)$ can be expanded around $\hat{\mathbf{x}}_0$ in Taylor series as:

$$\mathcal{P}(\mathbf{x}_0, t) = \mathcal{P}(\hat{\mathbf{x}}_0, t) + \frac{\partial \mathcal{P}}{\partial \mathbf{x}_0} \frac{1}{1!} (\mathbf{x}_0 - \hat{\mathbf{x}}_0) + \dots + \frac{\partial^n \mathcal{P}}{\partial^n \mathbf{x}_0} \frac{1}{n!} (\mathbf{x}_0 - \hat{\mathbf{x}}_0)^n \quad (3.8)$$

where, $\mathcal{P}(\hat{\mathbf{x}}_0, t)$ is the neighbouring trajectory to $\mathcal{P}(\mathbf{x}_0, t)$ and $\hat{\mathbf{x}}_0 \cong \mathbf{x}_0$. Neglecting all terms except for the first correction term in Equation: 3.8 for now, one may write:

$$\mathbf{Z} - \mathcal{P}(\mathbf{x}_0, t) = \mathbf{Z} - \mathcal{P}(\hat{\mathbf{x}}_0, t) - \mathbf{F}(\mathbf{x}_0 - \hat{\mathbf{x}}_0) \quad (3.9)$$

where,

$$\mathbf{F} = \left. \frac{\partial \mathcal{P}}{\partial \mathbf{x}_0} \right|_{\mathbf{x}_0 = \hat{\mathbf{x}}_0}$$

Substituting Equation: 3.9 in Equation: 3.6 one may write:

$$\begin{aligned} \mathbf{F}^T (\mathbf{Z} - \mathcal{P}(\mathbf{x}_0, t) - \mathbf{F} \Delta \hat{\mathbf{x}}_0) &= 0 \\ \mathbf{F}^T (\mathbf{Z} - \mathcal{P}(\mathbf{x}_0, t)) &= \mathbf{F}^T \mathbf{F} \Delta \hat{\mathbf{x}}_0 \\ \Delta \hat{\mathbf{x}}_0 &= (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T (\mathbf{Z} - \mathcal{P}(\mathbf{x}_0, t)) \end{aligned} \quad (3.10)$$

where, $\Delta \hat{\mathbf{x}}_0 = \mathbf{x}_0 - \hat{\mathbf{x}}_0$. The value of $\Delta \hat{\mathbf{x}}_0$ is still not the solution which minimizes the difference between the two trajectories \mathbf{Z} and \mathcal{P} due to the neglect and removed terms of higher order Taylor series expansion in Equation: 3.8. However, one may formulate an iterative scheme by repeatedly updating the value of $\hat{\mathbf{x}}_0$ in \mathcal{P} using newly computed $\Delta \hat{\mathbf{x}}_0$ from Equation: 3.10. The iterative scheme can be formulated as:

$$\hat{\mathbf{x}}_0^{j+1} = \hat{\mathbf{x}}_0^j + (\mathbf{F}^{jT} \mathbf{F}^j)^{-1} \mathbf{F}^{jT} (\mathbf{Z} - \mathcal{P}(\hat{\mathbf{x}}_0^j, t)) \quad (3.11)$$

where,

$$\mathbf{F}^j = \left. \frac{\partial \mathcal{P}}{\partial \mathbf{x}_0} \right|_{\mathbf{x}_0 = \hat{\mathbf{x}}_0^j}$$

The value of \mathbf{x}_0^{j+1} in Equation 3.11 will be repeated until the difference $\Delta\mathbf{x}_0^{j+1} = \mathbf{x}_0^{j+1} - \mathbf{x}_0^j$ is less than some selected tolerance or the variance σ^2 asymptotically reaches minimum. At this time $\hat{\mathbf{x}}_0$ has converged to define the optimum trajectory. The fundamental concept of this analysis methodology is shown in Figure: 3-2.

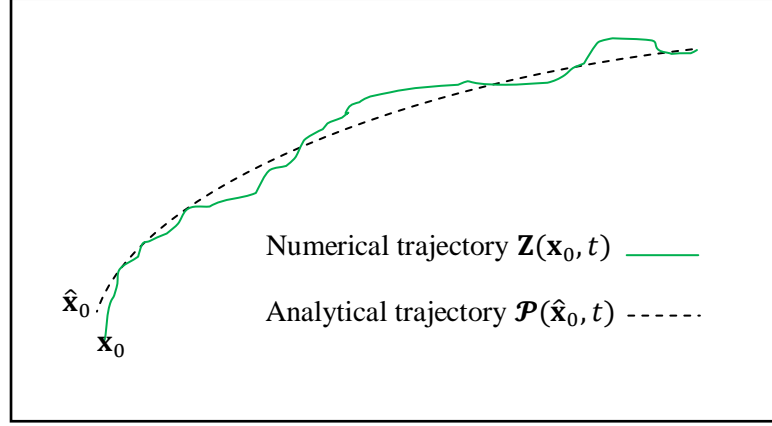


Figure 3-2: Concept of methodology for linearized orbital analysis. The numerical trajectory is obtained by numerical integration of equations of motion. Analytical trajectory is obtained from approximate analytic equations of motion using IC estimates $\hat{\mathbf{x}}_0$. The estimated IC $\hat{\mathbf{x}}_0$ minimizes the variance between Numerical trajectory and analytical trajectory.

3.3 Two Body Equation Review

The Newton's second law and his universal law of gravitation is essentially a starting point for any study of orbital motion, especially when combined with Kepler's law [83],[84]. Employing these laws the 2 *body* equation of motion can be derived as [12]:

$$\ddot{\vec{r}} = -\frac{\mu_E}{r^2} \frac{\vec{r}}{r} \quad (3.12)$$

where, \vec{r} is the position vector of a satellite in Earth Central Inertial (ECI) coordinates expressed as $\vec{r} = [X \ Y \ Z]^T$, $r = \sqrt{X^2 + Y^2 + Z^2}$. The value of the gravitational parameter μ_E in Equation: 3.12 is

computed as:

$$\mu_E = GM_E = 398600.4418 \text{ km}^3 / \text{sec}^2$$

$$G = 6.673 \times 10^{-20} \text{ km}^3 / \text{kg} / \text{sec}^2, \text{ Earth Gravitational constant}$$

$$M_E = 5.973 \times 10^{24} \text{ kg}, \text{ Mass of Earth}$$

The ECI coordinate frame is defined such that X axis points to the vernal point in the equatorial plane of the Earth, the Z axis is the axis of rotation of the Earth in positive direction, and Y is defined by the right-hand rule (see Figure: 3-3).

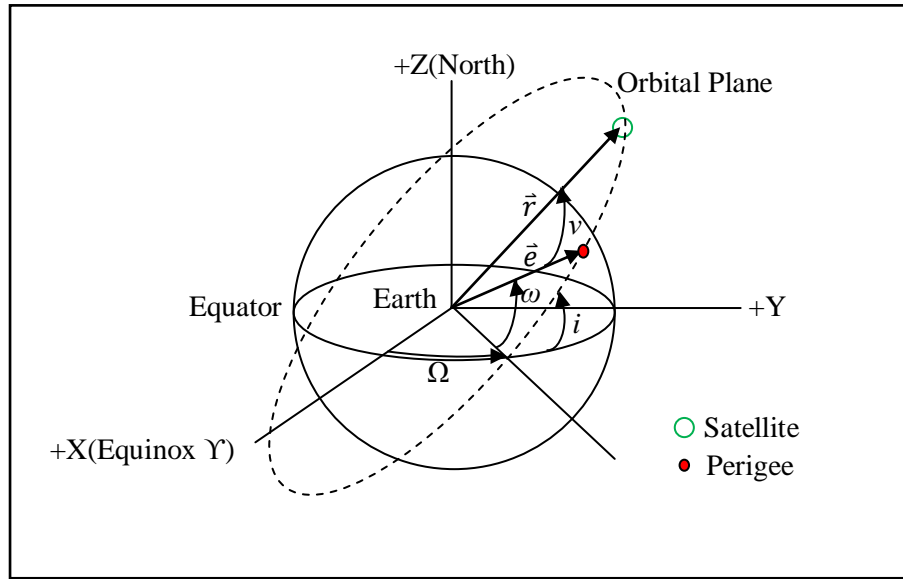


Figure 3-3: Earth Central Inertial (ECI) Coordinates frame. X -axis points towards Vernal point in equatorial plane of Earth (Equinox Y), Z -axis points towards North pole and Y -axis completes the right hand triad. Angular quantities of orbital elements, i = inclination, Ω = right ascension of the ascending node, ω = argument of perigee and v = true anomaly are also illustrated.

Alternatively, the acceleration for a 2 *body* problem in a spherical symmetric gravitational potential function can be expressed as gradient of the potential function expressed as:

$$\ddot{\vec{r}} = \nabla \Phi \quad (3.13)$$

$$\Phi = \frac{\mu_E}{r} \quad (3.14)$$

where, Φ is the gravitational potential function.

Six quantities define the state of a satellite in space. These quantities can be expressed in many equivalent forms. Whatever the form is, the collection of these quantities is called either a state vector $\mathbf{x} = [X \ Y \ Z \ \dot{X} \ \dot{Y} \ \dot{Z}]^T$ usually associated with position and velocity vectors, or an element set, normally used with scalar magnitudes and angular representations of the orbit called orbital elements. The six classical orbital elements are; a is the semi major axis, e is the eccentricity, i is the inclination, Ω is the Right Ascension of the Ascending Node (RAAN), ω is the argument of perigee and v is the true anomaly. Figure: 3-3 depicts angular quantities of classical orbital elements [12]. The definition of the semi major axis and eccentricity are [12]:

$$a = -\frac{\mu_E}{2\varepsilon} \quad (3.15)$$

$$\vec{e} = \frac{\left(v^2 - \frac{\mu_E}{r}\right)\vec{r} - (\vec{r} \cdot \vec{v})\vec{v}}{\mu_E} \quad (3.16)$$

$$\varepsilon = \frac{v^2}{2} - \frac{\mu_E}{r} \quad (3.17)$$

where, \vec{v} is the velocity vector in ECI coordinates expressed as $\vec{v} = [\dot{X} \ \dot{Y} \ \dot{Z}]^T$, $v = |\vec{v}|$, ε = orbital energy, and \vec{e} = eccentricity vector ($e = |\vec{e}|$) pointing towards the perigee (see Figure: 3-3). Under an axially symmetric gravitational potential the orbital energy (ε) remains constant [2]. A solution for Equation: 3.12 can be obtained using numerical integration techniques such as Runge-Kutta method i.e., RK-4 [13] by providing an initial state vector $\vec{x}_0 = [\vec{r}_0 \ \vec{v}_0]$ to the numerical algorithm.

3.3.1 Kepler's Equation

A solution of 2 *body* equation can be also be obtained analytically by solving the Kepler's equation [13]. The Kepler's equation allows to determine the relation of time and angular displacement of a satellite within an orbit. The Kepler's equation is mathematically expressed as:

$$E(t) - e \sin E(t) = n(t - t_p) \quad (3.18)$$

$$M = n(t - t_p) \quad (3.19)$$

where, M is the mean anomaly, $E(t)$ is the eccentric anomaly, e is the eccentricity, $n = \sqrt{\mu_E/a^3}$ (mean motion) and t_p is the time of perigee passage. The Kepler's equation relates the time t to the coordinates \hat{X} and \hat{Y} in the orbital plane of a satellite via the eccentric anomaly. A geometrical description of these quantities is described in Figure: 3-4.

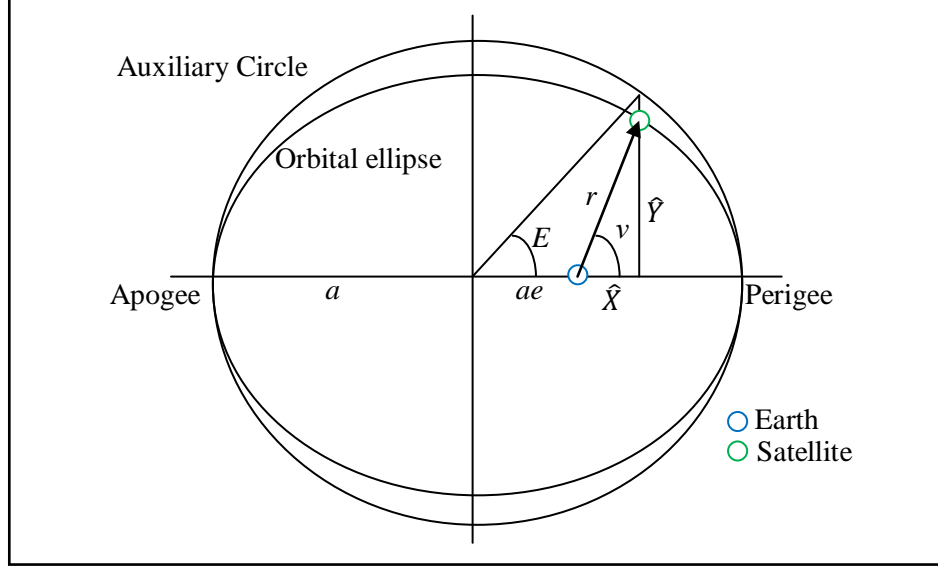


Figure 3-4: Orbital geometry for Kepler's equation defining the eccentric anomaly (E), true anomaly (v) and coordinates in orbital ellipse (plane) $[\hat{X} \ \hat{Y}]^T$. In the orbital plane, perigee is the point nearest to the centre of gravitational attraction and apogee is the point farthest.

The orbital coordinates in terms of Eccentric anomaly are expressed as:

$$\hat{X} = r \cos v =: a(\cos E - e) \quad (3.20)$$

$$\hat{Y} = r \sin v =: a\sqrt{1 - e^2} \sin E \quad (3.21)$$

where, $r = a(1 - e \cos E)$

One must know the time of perigee passage and the semi major axis in order to calculate the mean anomaly. Then one may find values of E that satisfy Equation: 3.18 and finally obtain \hat{X} and \hat{Y} in the orbital plane of satellite trajectory described by Equation: 3.20-3.21. The Kepler's equation is usually solved using iterative methods like Newton-Raphson [100]. The solution of Kepler's equation is found out in a perifocal coordinate system (described in next section). A more useful representation of an orbit for our analysis is in ECI coordinate system, as the equations of motion (Equation: 3.12) are expressed in that coordinate system.

3.3.2 Conversion from Perifocal to ECI Coordinates

In order to represent the position and velocity of a satellite in ECI coordinates, first the satellite's three dimensional position and velocity are expressed in perifocal coordinate system. A perifocal coordinate

system is described as [14]:

$$\mathbf{P} = \frac{\vec{e}}{|\mathbf{e}|}, \quad \mathbf{Q} = \mathbf{W} \times \mathbf{P}, \quad \mathbf{W} = \frac{\vec{r} \times \vec{v}}{|\vec{r} \times \vec{v}|} \quad (3.22)$$

Using perifocal coordinates one may express the three dimensional position by [13]:

$$\vec{r}_{perifocal} = \hat{X}\mathbf{P} + \hat{Y}\mathbf{Q} \quad (3.23)$$

$$= a(\cos E - e)\mathbf{P} + a\sqrt{1 - e^2} \sin E \mathbf{Q} \quad (3.24)$$

and the velocity by

$$\vec{v}_{perifocal} = \dot{\hat{X}}\mathbf{P} + \dot{\hat{Y}}\mathbf{Q} \quad (3.25)$$

$$= \frac{\sqrt{GM_E a}}{r} \left(-\sin E \mathbf{P} + \sqrt{1 - e^2} \cos E \mathbf{Q} \right) \quad (3.26)$$

The classical orbital angular elements of (i, Ω, ω) are employed in rotation transformation to convert perifocal coordinate system into ECI. The rotation matrix is given as:

$$\mathbf{PQW} = \mathfrak{R}_z(-\Omega) * \mathfrak{R}_x(-i) * \mathfrak{R}_z(-\omega) \quad (3.27)$$

where, $\mathfrak{R}_z(.)$ and $\mathfrak{R}_x(.)$ are rotation matrices for rotation about “z” and “x” axes respectively. For example for θ being the rotation angle the individual rotation matrices can be computed as:

$$\mathfrak{R}_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & +\cos\theta & +\sin\theta \\ 0 & -\sin\theta & +\cos\theta \end{bmatrix}, \quad \mathfrak{R}_z(\theta) = \begin{bmatrix} +\cos\theta & +\sin\theta & 0 \\ -\sin\theta & +\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.28)$$

Finally, the three dimensional ECI position and velocity are obtained respectively as [13]:

$$\begin{aligned} \vec{r} &= \mathbf{PQW} * \vec{r}_{perifocal} \\ \vec{v} &= \mathbf{PQW} * \vec{v}_{perifocal} \end{aligned}$$

(3.29)

The above (Equation 3.29) provides an analytic solution to the unperturbed 2 *body* equation. However, it requires a solution of the Kepler's equation at each time step using iterative methods which may develop some convergence problems [13].

3.4 Perturbation Due to Oblate Earth – J_2

The gradient of the gravitational potential function for a spherical Earth model will yield acceleration as expressed in Equation: 3.12. In reality the Earth is closer to an oblate spheroid therefore one may now consider perturbing forces due to non-spherical gravitational potential function. A non-spherical potential function which is symmetric about Earth polar axis could be expressed as [12]:

$$\mathfrak{U} = \frac{\mu_E}{r} \left[1 - \sum_l J_l \left(\frac{R_E}{r} \right)^l P_l[\sin \phi] \right] \quad (3.30)$$

where, J_l is the coefficients of zonal spherical harmonic representing the shape of Earth, $P_l[.]$ is the Legendre polynomials of degree l , ϕ is the geocentric latitude of satellite (see Figure: 3-5) and R_E is the Equatorial radius of Earth ($R_E = 6378.137 \text{ km}$).

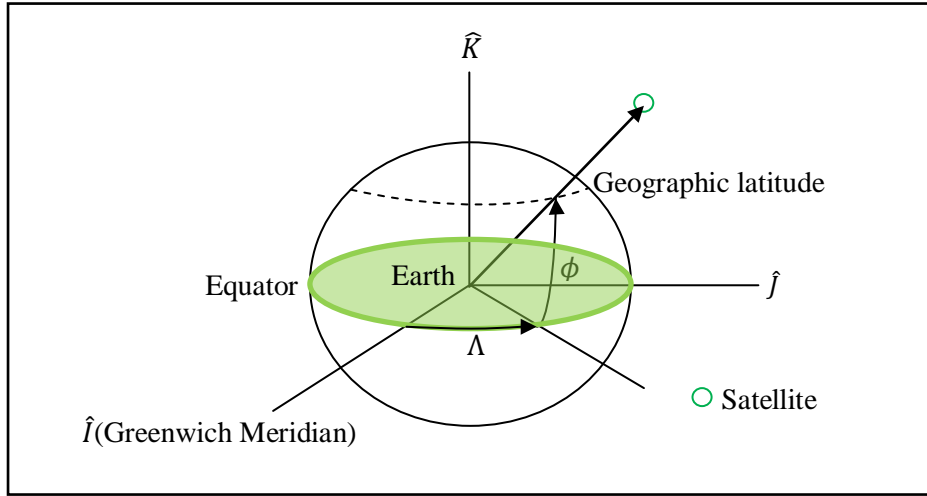


Figure 3-5: Geometrical description of geocentric latitude (ϕ) and longitude (Λ) of satellite in coordinate system fixed to Earth known as Earth Centred Earth Fixed (ECEF). Note oblate shape around Earth Equator (shown green) which is responsible for J_2 perturbation.

The expansion of Equation: 3.30 in terms of Legendre polynomials $P_l[\sin \phi]$ for order up to $l = 4$ is expressed as [8]:

$$\begin{aligned} \mathfrak{U} = \frac{\mu_E}{r} & \left[1 - J_2 \left(\frac{R_E}{r} \right)^2 \left(\frac{1}{2} \{ 3 \sin^2 \phi - 1 \} \right) \right. \\ & - J_3 \left(\frac{R_E}{r} \right)^3 \left(\frac{1}{2} \{ 5 \sin^3 \phi - 3 \sin \phi \} \right) \\ & \left. - J_4 \left(\frac{R_E}{r} \right)^4 \left(\frac{1}{8} \{ 35 \sin^4 \phi - 30 \sin^2 \phi + 3 \} \right) \right] \end{aligned} \quad (3.31)$$

Similarly to Equation: 3.13 the acceleration due to gravity can be derived by taking the gradient of this potential function (Equation: 3.31). The acceleration for terms up to $l = 2$ can be expressed as [13],[12],[8]:

$$\begin{aligned} \ddot{\vec{r}} &= -\frac{\mu_E}{r^2} \frac{\vec{r}}{r} + a_G \\ a_{GX} &= -\frac{3}{2} \frac{\mu_E J_2 R_E^2}{r^4} \left(1 - \frac{3Z^2}{r^2} \right) \frac{X}{r} \\ a_{GY} &= -\frac{3}{2} \frac{\mu_E J_2 R_E^2}{r^4} \left(1 - \frac{3Z^2}{r^2} \right) \frac{Y}{r} \\ a_{GZ} &= -\frac{3}{2} \frac{\mu_E J_2 R_E^2}{r^4} \left(3 - \frac{3Z^2}{r^2} \right) \frac{Z}{r} \end{aligned} \quad (3.32)$$

where, the vector $[a_{GX} \ a_{GY} \ a_{GZ}]^T$ contains the components of acceleration in ECI coordinates due to the second spherical harmonic, $J_2 = 1.08262668355 \times 10^{-3}$. Numerical integration techniques such as RK-4 algorithm may be utilized to obtain a solution for Equations: 3.32. In general, and analytic solution for non-spherical Earth 2 *body* equation of motion is obtained by linearization and approximation on acceleration terms briefly shown in Table: 3-1 [2],[15],[16],[93]. Therefore, the two orbital descriptions (numerical and analytical) are not exactly identical. The second zonal harmonic term J_2 has small perturbative acceleration compared to main spherical gravity (see Table: 3-1) on satellite orbits. However, these orbits are characterized by secular and periodic changes in their orbital elements. Secular variations in a particular element change linearly over time and cause unbounded error growth. Periodic changes are of two types: (1) short periodic, and (2) long periodic, depending upon the amount of time required for the effect to repeat. Short periodic effects repeat on the order of the satellite period or even can be less frequent. Long periodic effects have cycles significantly larger than one orbital period, usually one or two times longer than short periodic [12],[2]. The argument of perigee, right ascension of the ascending node and true anomaly have secular variations which grow over time (see Figure: 3-11). These elements and other remaining elements i.e., (a, e, i) have both short periodic and long periodic variations [15],[16] (see Figure: 3-9). There are no long periodic variations in the orbit due to J_2 perturbation [2]. Contrary to J_2 perturbed orbits, the orbital elements

for a 2 *body* equation without J_2 , as expressed in Equation: 3.12 would remain constant except for the true anomaly which continuously changes (see Figure: 3-8 and 3-10). One may visually inspect the effect of secular and periodic variations in the J_2 perturbed orbits compared without these effects as shown in Figure: 3-6 and 3-7 in ECI coordinates. Figures: 3-6 to 3-11 are simulated for a typical LEO satellite with $a = 7000$ km, $e = 0.0001$, $i = 98$ deg, $\Omega = 10$ deg, $\omega = 10$ deg, and $\nu = 20$ deg.

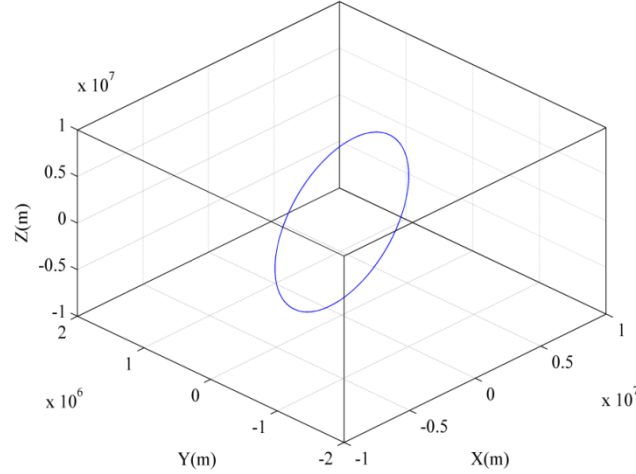


Figure 3-6: Time history of a satellite orbit in ECI coordinates obtained from numerical integration of equations of motion (Equation: 3.12) for one day. Notice a near circular motion without any periodic or secular changes in the orbit.

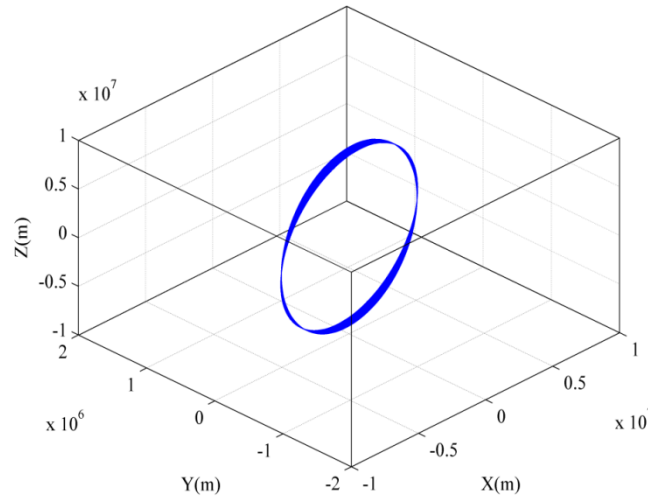


Figure 3-7: Time history of a satellite orbit in ECI coordinates obtained from numerical integration of equations of motion (Equation: 3.32) for one day. Notice variations in the orbit due to periodic and secular (drift) effects of J_2 perturbation.

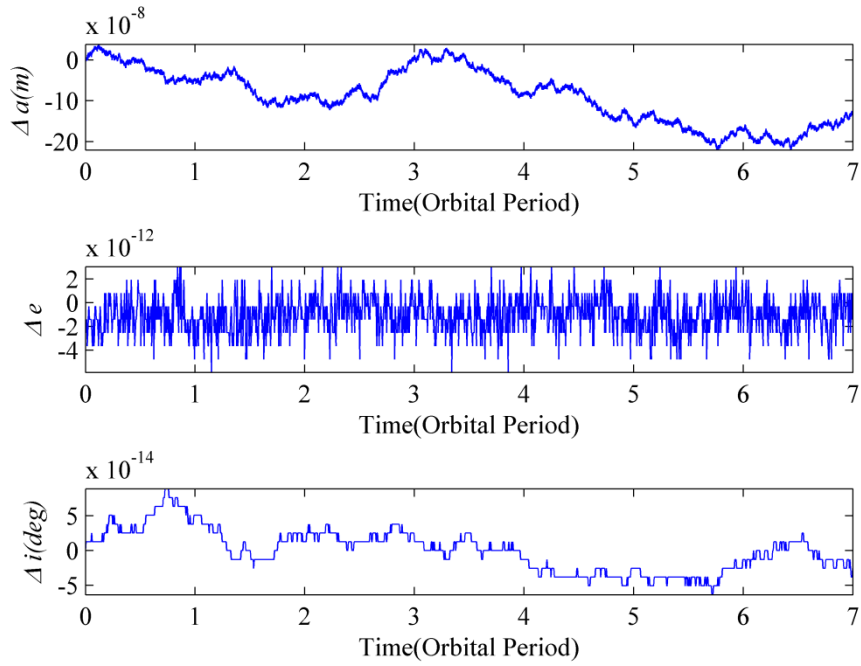


Figure 3-8: Time history of variations (Δ) in orbital elements of a (top), e (middle), and i (bottom) for a 2 *body* equation (Equation: 3.12) for 7 orbital periods (approximately half a day) for LEO satellite. Notice these elements remain almost constant under a spherically symmetric geopotential.

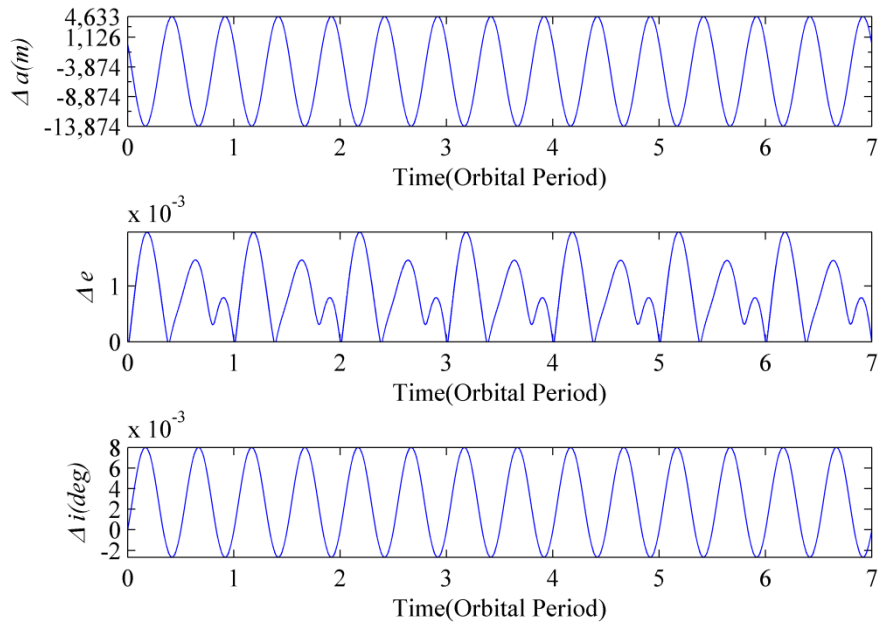


Figure 3-9: Time history of variations (Δ) in orbital elements of a (top), e (middle), and i (bottom) for J_2 perturbed 2 *body* equation (Equation: 3.32). Notice periodic variations in the orbit under non-spherical geopotential. However, these elements do not have secular (drift) effects.

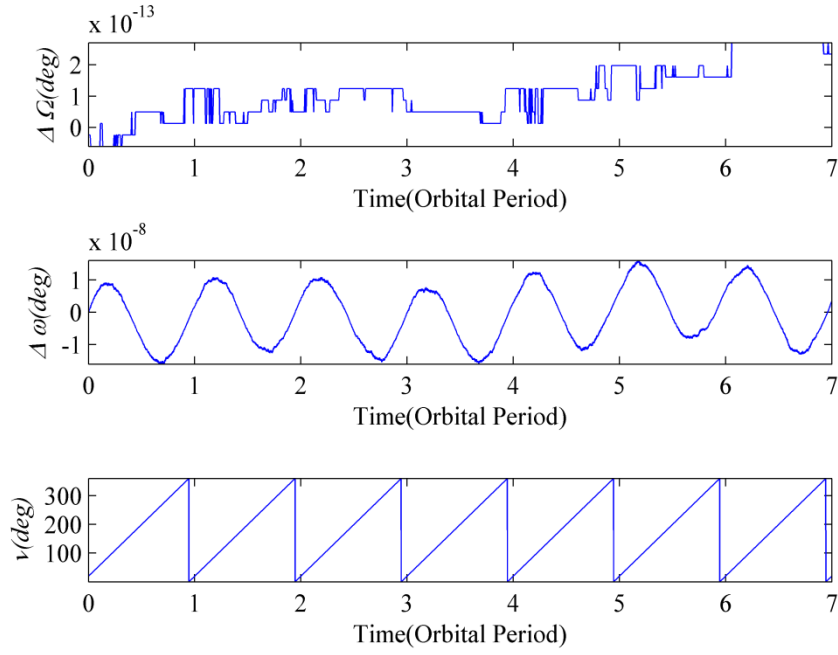


Figure 3-10: Time history of variations (Δ) in angular quantities of orbital elements Ω (top) and ω (middle) for a 2 *body* equation under a spherically symmetric geopotential (Equation: 3.12). These elements remain almost constant except for v (bottom) which changes by 360 degrees over one orbital period.

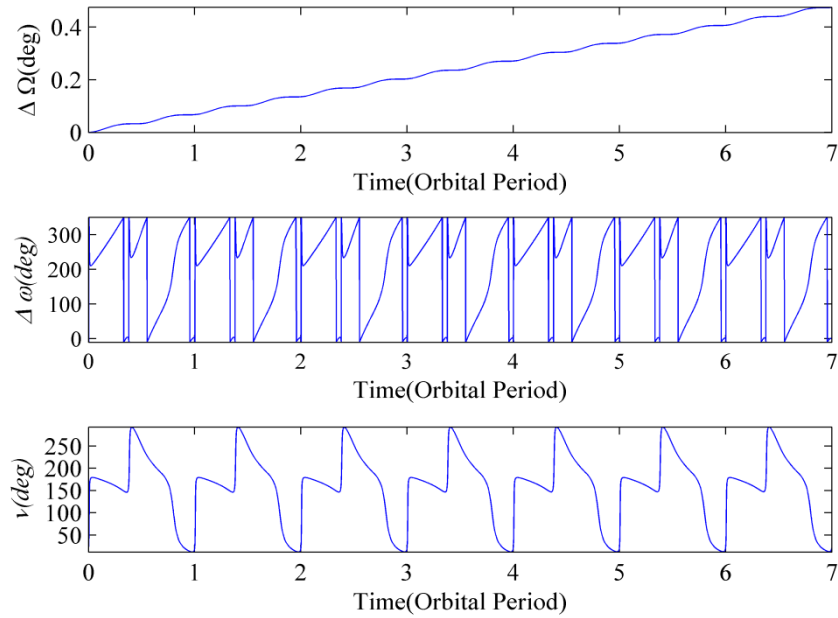


Figure 3-11: Time history of variations (Δ) in angular quantities of orbital elements Ω (top), ω (middle) for a J_2 perturbed 2 *body* equation (Equation: 3.32). The v (bottom) varies between 0 to 360 deg over an orbital period. Notice the small periodic oscillation and significant linear secular / drift variations in these elements under a non-spherical geopotential.

3.5 Analysis of Absolute Satellite Orbital Dynamics

Satellite absolute dynamics are referred here as orbital motion around a central gravitational field such as the Earth. Consider the unperturbed 2 *body* equation of motion (Equation: 3.12) and dynamic model due to non-spherical Earth gravitational potential with J_2 perturbation (Equation: 3.32) as the full nonlinear orbital models for the analysis. Therefore, in this section the following two analytic models will be considered for comparison with these nonlinear orbital models:

- A solution of Kepler's equation for unperturbed orbit (Section: 3.3.1 and 3.3.2) [13].
- The Epicyclic model for oblate Earth by Hashida and Palmer (explained later in the chapter) [2].

The equations of motion used in this analysis are expressed in ECI coordinate system. Therefore, errors between the numerical and analytical trajectory in these coordinates would be required for the evaluation. However, more useful comparison of errors from the point of view of visualization can be done, by transforming errors from ECI coordinates into a Local Vertical Local Horizontal (LVLH) coordinate system. The LVLH coordinate system is also known as satellite coordinate system. The system moves with the satellite and its origin is the centre of gravity of the satellite (see Figure: 3-12).

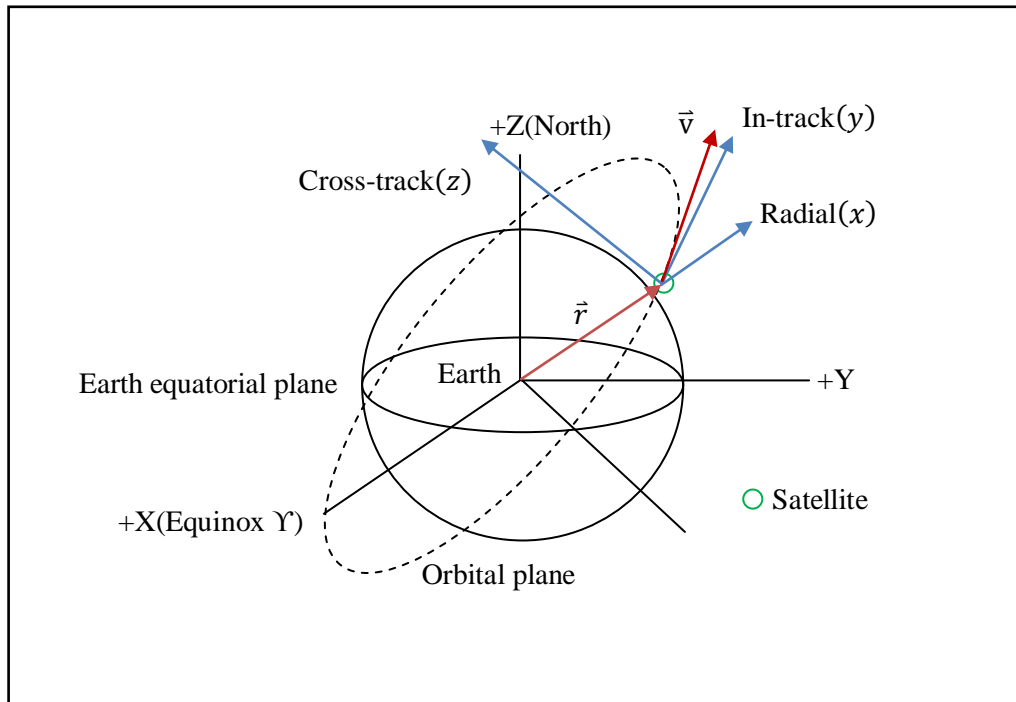


Figure 3-12: Illustration of the Local Vertical Local Horizontal (LVLH) system centred at satellite centre of gravity. (x) axis points from the Earth's centre along the radius vector towards the satellite as it moves along the orbit. (y) axis points in the direction of velocity vector (not necessarily parallel) and is perpendicular to radius vector. The (z) axis is normal to the orbital plane.

The reference plane is the orbital plane of the satellite, and the principle direction is the radius vector from the centre of the Earth to the satellite. The x -axis points from the centre of the Earth along the radius vector towards the satellite, as it moves through the orbit. This motion is referred as *radial* direction. The z -axis is fixed along the direction normal to the orbital plane and is termed as *cross-track* direction. The y -axis is perpendicular to the radius vector and is not aligned with the velocity vector except for Keplerian orbits and elliptical orbits at *perigee* and *apogee* (see Figure: 3-4 and 3-12 for illustration) [14]. It is referred as in-track direction.

In the following analysis the term reference orbit will be used for numerically obtained trajectory and analytical trajectory for a trajectory obtained from an analytic model. The IC at epoch time (t_0) in terms of classical orbital elements $\mathbf{y}(t_0)$ selected for the reference LEO orbit are:

$$\begin{aligned} a &= 6863.100 \text{ km} \\ e &= 0.0001 \\ I &= 98 \text{ deg} \\ \Omega &= 0 \text{ deg} \\ \omega &= 0 \text{ deg} \\ \nu &= 0 \text{ deg} \\ M_0 &= 0 \text{ deg} \end{aligned} \tag{3.33}$$

where, M_0 is the mean anomaly at epoch.

3.5.1 Analysis of Kepler's Equation

The Kepler's equation provides an analytical solution for an unperturbed orbit under a spherically symmetric geopotential. It is expected that trajectories determined numerically and analytically for this problem would be sufficiently close to each. Moreover, as seen in Section: 3.4, the acceleration due to J_2 produces significant variations in the orbit in terms of secular and periodic effects. Therefore, one would like to evaluate the analytic solution of the Kepler's equation with J_2 perturbed nonlinear equations of motion as well. Thus in this section an analysis would be carried out on following:

- Analytic and numerical solutions of unperturbed 2 *body* equation.
- Analytic solution of the Kepler's equation compared with J_2 perturbed 2 *body* equation.

3.5.1.1 Unperturbed Two Body Equation and Analytic Solution of Kepler's Equation

In this section the analytic solution of the Kepler's equation given in Equation: 3.29 and numerical solution of 2 *body* equation (Equation: 3.12) are being compared. The initial classical orbital elements given in Equation: 3.33 are firstly converted into ECI position and velocity vector using transformation routines provided in Appendix-A [12]. Given IC $\mathbf{x}_0 = [\vec{r}_0, \vec{v}_0]^T$ in terms of ECI coordinates, numerical integration of Equation: 3.12 is computed using RK-4 with a step size

($\Delta t = 5$ sec) for a time of 10 orbital periods (3/4 of a day for LEO satellites). This would form a large column vector $\mathbf{Z}(\mathbf{x}_0, t)$ of $(6k)$ components. The analytical propagation of orbit trajectory in ECI coordinates is obtained by solving the Kepler's Equation: 3.18 from initial orbital elements (Equation: 3.33). This would form a column vector $\mathcal{P}(\mathbf{x}_0, t)$ of $(6k)$ components. In order to utilize the estimation algorithm (Equation: 3.11) an analytic expression for matrix (\mathbf{F}) is needed which could be expressed as [13]:

$$\frac{\partial \mathcal{P}(\mathbf{x}_0, t)}{\partial \mathbf{x}_0} = \frac{\partial \mathcal{P}(\mathbf{x}_0, t)}{\partial \mathbf{y}(t)} \cdot \frac{\partial \mathbf{y}(t)}{\partial \mathbf{y}(t_0)} \cdot \left(\frac{\partial \mathbf{x}_0}{\partial \mathbf{y}(t_0)} \right)^{-1} \quad (3.34)$$

where, \mathbf{y} is the $(a, e, i, \Omega, \omega, M)^T$, t_0 is the epoch time, \mathbf{x}_0 is the IC in ECI coordinates,

$\frac{\partial \mathcal{P}(\mathbf{x}_0, t)}{\partial \mathbf{y}(t)}$ is the partial derivative of satellite trajectory with respect to orbital elements at $\mathbf{y}(t)$

$\frac{\partial \mathbf{y}(t)}{\partial \mathbf{y}(t_0)}$ is the partial derivative of orbital elements with respect to orbital elements at epoch $\mathbf{y}(t_0)$

$\left(\frac{\partial \mathbf{x}_0}{\partial \mathbf{y}(t_0)} \right)^{-1}$ is the inverse of partial derivative of epoch state vector \mathbf{x}_0 with respect to orbital elements at epoch $\mathbf{y}(t_0)$. The partial derivatives expressed in Equation: 3.34 are obtained from Ref. [13] (see Appendix-B). The comparison with a reference trajectory is expressed as relative position and velocity deviations in LVLH coordinate frame of reference satellite. As shown in Figures: 3-13 and 3-14 the positional errors are in order of 10^{-4} m and velocity errors in order 10^{-8} m/s. This shows that reference nonlinear model and analytical solution are in close conformity to each other. The comparison of estimated IC for the analytical trajectory and the chosen IC (Equation: 3.33) for nonlinear trajectory is given in Table: 3-2.

Orbital Elements	IC of Numerical Trajectory \mathbf{x}_0	IC of Analytical Trajectory (Output of estimator - $\hat{\mathbf{x}}_0$)
a	6863.100 km	6863.099 km
e	0.0001	9.9999×10^{-5}
i	98 deg	~ 98 deg
Ω	0 deg	0 deg
ω	0 deg	0 deg
M_0	0 deg	1.4787×10^{-9}

Table 3-2: Comparison of IC for Numerical and Analytical trajectories for unperturbed Kepler's Equation. The difference in IC for (i) is of the order of (10^{-16}).

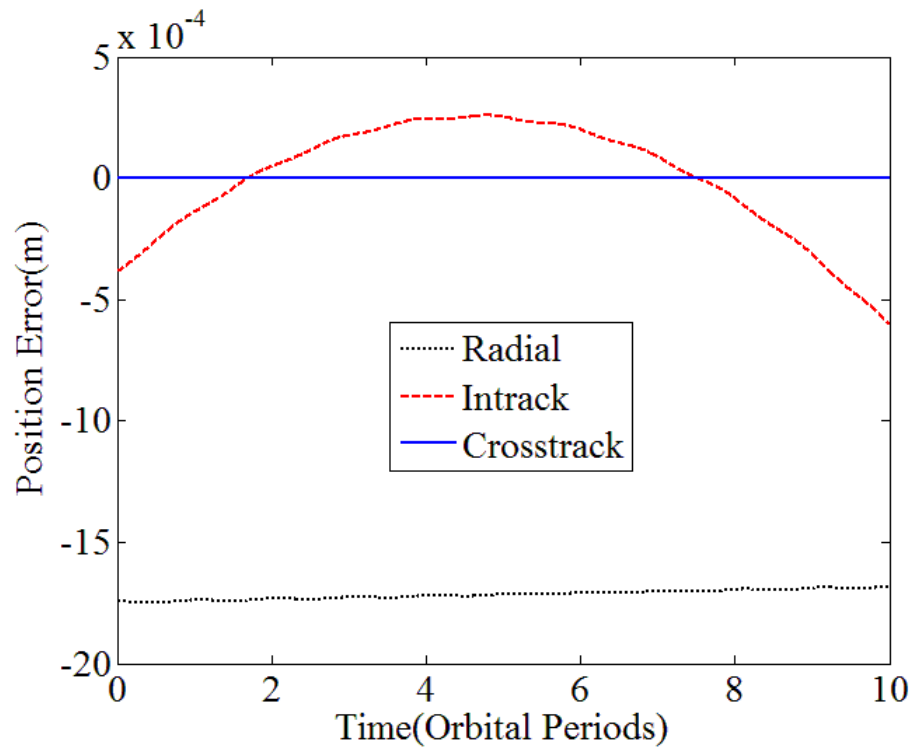


Figure 3-13: Time history of position errors for analytic solution of Kepler's equation compared with numerical trajectory of unperturbed 2 *body* equation.

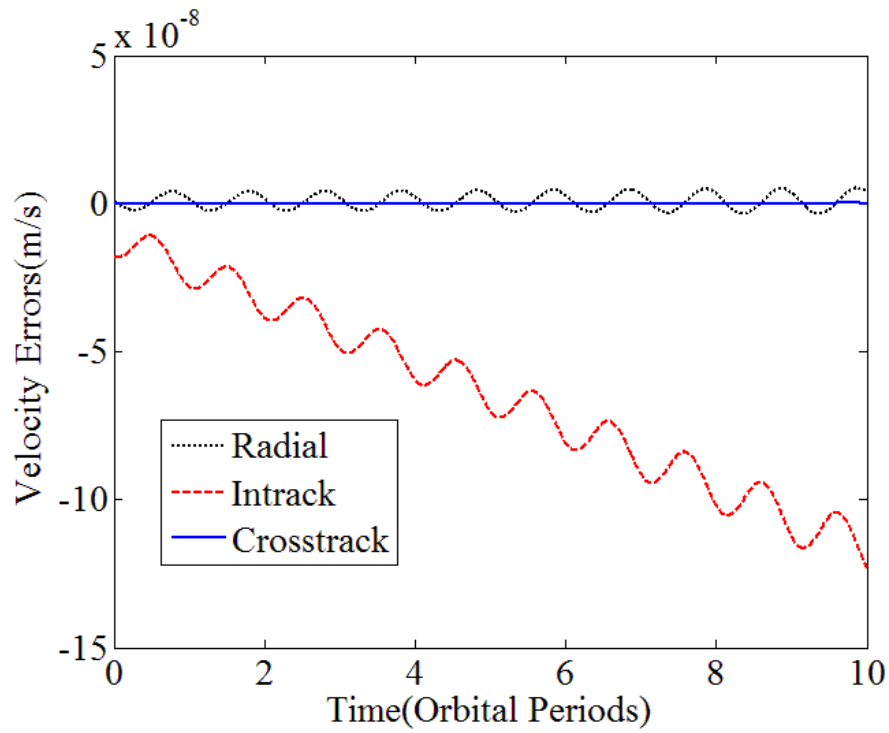


Figure 3-14: Time history of velocity errors for analytic solution of Kepler's equation compared with numerical trajectory of unperturbed 2 *body* equation.

3.5.1.2 J_2 Perturbed Two Body Equation and Analytic Solution of Kepler's Equation

The comparison of the analytic solution of the Kepler's equation with J_2 perturbed 2 *body* equation is carried out in this section. Similarly the previous analysis for acceleration under spherically symmetric geopotential, now J_2 perturbed 2 *body* Equation: 3.32 would be integrated numerically using RK-4. The orbital elements given in Equation: 3.33 are being used as IC for numerical integration in this analysis as well. The methodology for estimating IC for analytical solution remains the same as done in Section: 3.5.1.1. The comparison of estimated IC for the analytical trajectory is presented in Table: 3-3.

Orbital Elements	IC of Numerical Trajectory \mathbf{x}_0	IC of Analytical Trajectory (Output of estimator - $\hat{\mathbf{x}}_0$)
a	6863.100 km	6859.714 km
e	0.0001	3.84×10^{-4}
i	98 deg	98.016 deg
Ω	0 deg	0.353 deg
ω	0 deg	178.85 deg
M_0	0 deg	181.195 deg

Table 3-3: Comparison of IC for numerical and analytical trajectory for J_2 perturbed 2 *body* equation once compared with analytic solution of Kepler's equation.

The time history of the position and velocity in LVLH coordinate frame of the reference satellite are shown in Figures: 3-15 and 3-16. The maximum positional and velocity errors are summarized in Table: 3-4. As anticipated, there are significant errors in all three directions due to neglecting of J_2 acceleration in the Kepler's Equation. The main cause of errors stems from secular variations in a J_2 perturbed orbit. As shown in Figure: 3.15 and 3.16, the worst case error is observed in cross-track direction. The cross track motion is primarily due to a difference in inclination and RAAN. Therefore, the secular growth in RAAN and periodic variation in the inclination (see Figure: 3-9 and 3-11) are responsible for these errors. Radial errors have significant deviations due to periodic terms in the semi-major axis and eccentricity (see Figure: 3-9), whereas, in-track errors are due to secular drift in argument of perigee and mean anomaly (see Figure: 3-11).

However, all these errors are considerably less if compared with orbits once the IC are not estimated. This means that the IC (given in Equation: 3.33) is used to generate both analytical and numerical trajectories. The plots of positional and velocity errors without estimated IC are shown in Figure: 3-17 and 3-18. A significant drift term in in-track motion is due to unbounded error growth in the secular

term in argument of perigee and mean anomaly. The effect is more pronounced due to a difference in the mean motion of the satellite caused by variation of the semi-major axis. The radial errors are more or less periodic in nature, with a constant offset term which is due to the offset term and short periodic variations in osculating semi-major axis and eccentricity (see Figure: 3-9).

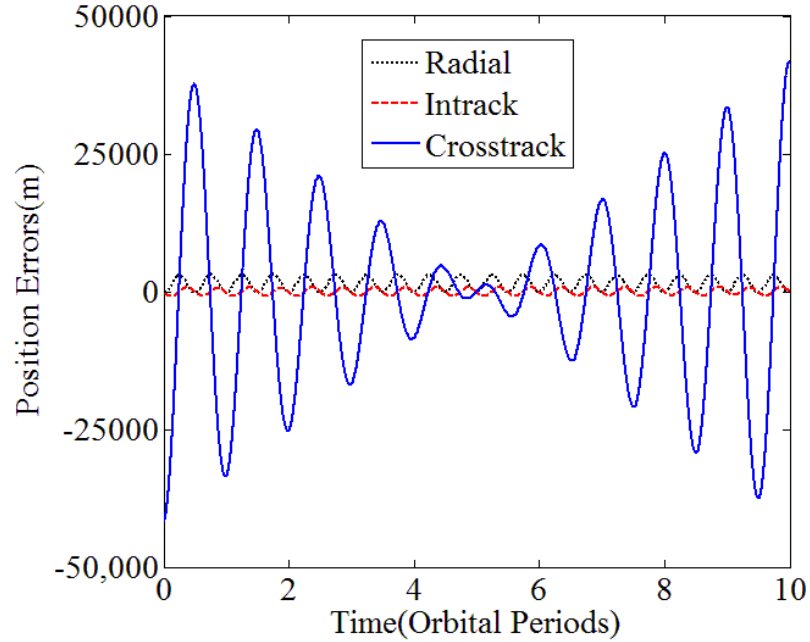


Figure 3-15: Time history of position errors for analytic solution of the equation of Kepler compared with numerical trajectory of J_2 perturbed 2 *body* equation.

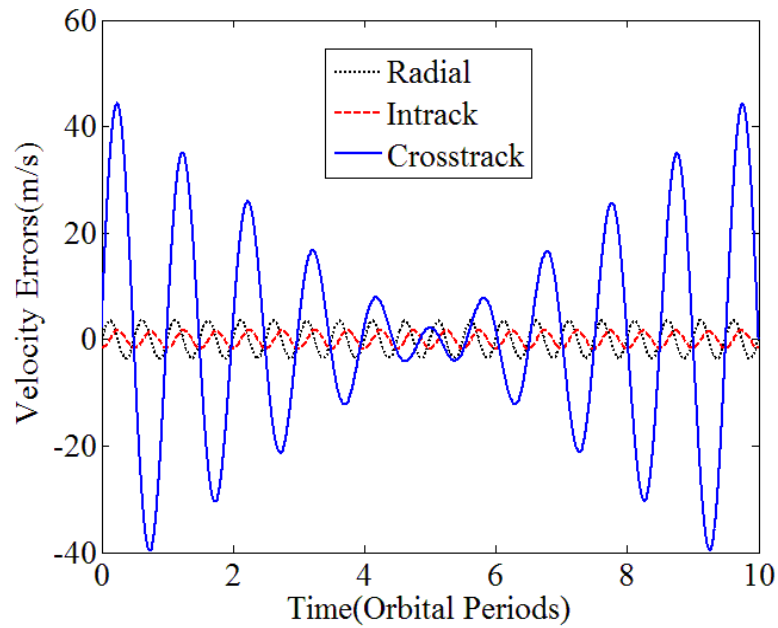


Figure 3-16: Time history of velocity errors for analytic solution of Kepler's equation compared with numerical trajectory of J_2 perturbed 2 *body* equation.

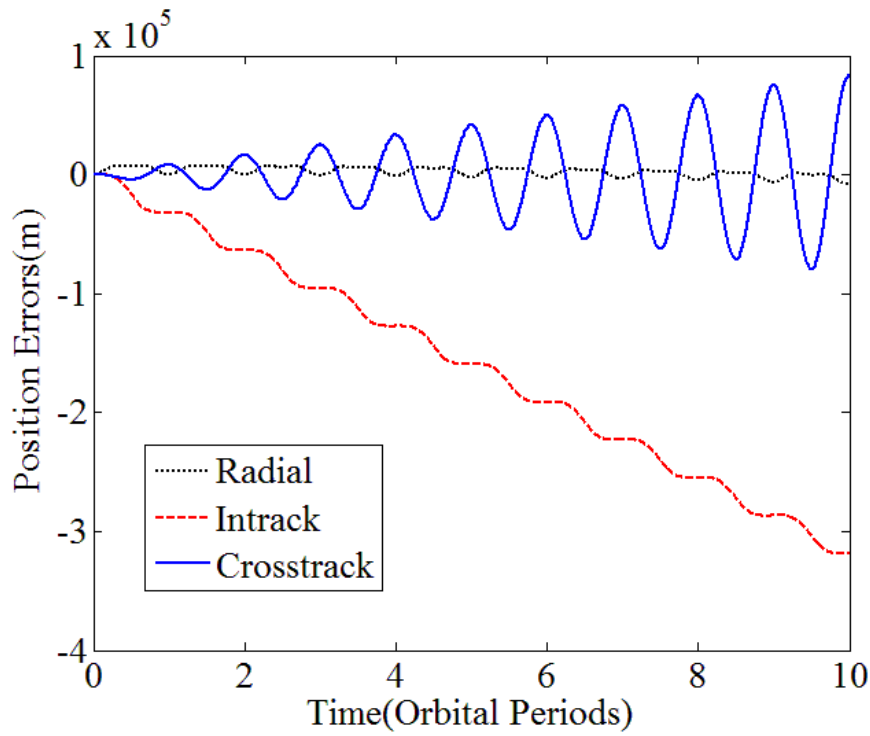


Figure 3-17: Time history of position errors for analytic solution of Kepler's equation compared with numerical trajectory of J_2 perturbed 2 body equation without estimating IC.

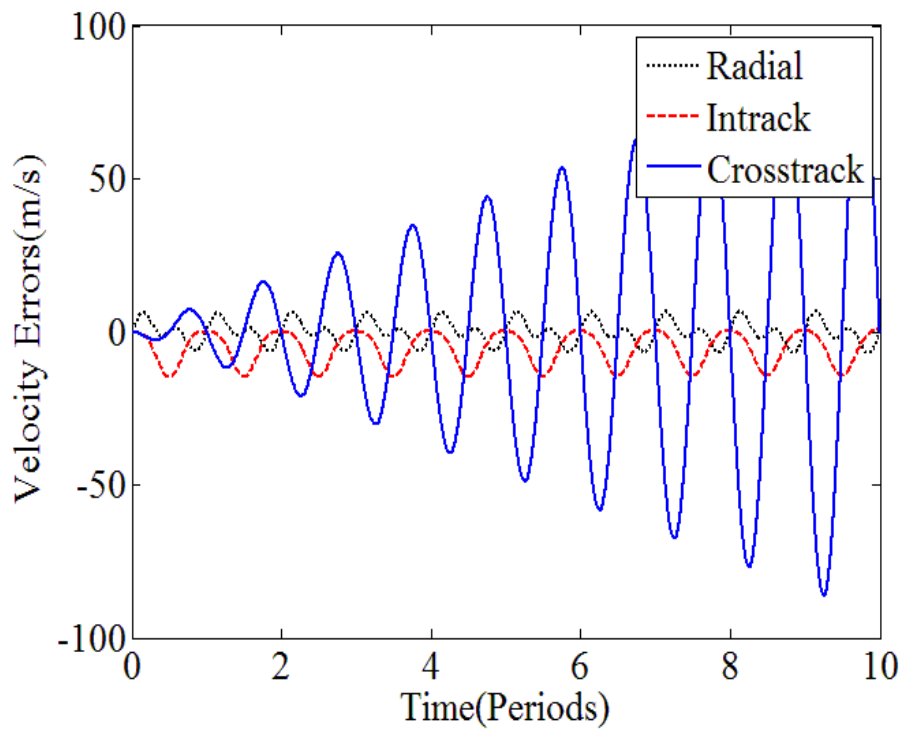


Figure 3-18: Time history of velocity errors for analytic solution of Kepler's equation compared with numerical trajectory of J_2 perturbed 2 body equation without estimating IC.

IC used for propagation of Kepler's Equation	Position Errors (m)			Velocity Errors (m/s)		
	R	I	C	R	I	C
\mathbf{x}_0	7906	318834	83668	6.82	14.78	90.88
$\hat{\mathbf{x}}_0$	3136	852	41904	3.65	1.77	44.38
R – Radial, I – In-track, C – Cross-track $\hat{\mathbf{x}}_0$ – Output of the estimator \mathbf{x}_0 – IC of Numerical trajectory						

Table 3-4: Summary of the Maximum Absolute Position and Velocity Errors in LVLH Coordinates over 10 Orbital Periods for analytic solution of Kepler's Equation compared with the numerical solution of J_2 perturbed 2 *body* Equation: 3.32.

3.5.2 Epicyclic Motion of Satellite about an Oblate Planet

An analytic formulation for a near circular epicyclic orbit of a satellite around an oblate Earth by Hashida and Palmer [2] is now being considered for analysis. The model has a simple analytic form, describing all the geopotential terms arising from the Earth zonal harmonics. In this analysis terms up to J_2 are utilized. The state of a satellite in an epicyclic orbit is defined by a set of six osculating (instantaneous) spherical coordinates expressed in ECI (Equation: 3.35). The position of the satellite is described by a redundant set of four coordinates (r, λ, I, Ω) and velocity by (v_r, v_θ) . The pictorial representation of these coordinates is shown in Figure: 3-19. Inclination (I) and right ascension of the ascending node (Ω) defines the orbital plane of a satellite (the plane containing the position and velocity vectors), and radial coordinate (r) and argument of latitude (λ) locate the position of the satellite on that plane. The argument of latitude (λ) is analogous to sum of argument of perigee (ω) and true anomaly (ν) for circular orbits (see Figure: 3.3). However, it is measured from the time when the satellite crosses the initial ascending node while travelling from the southern hemisphere to the northern hemisphere. The components of velocity are radial velocity (v_r) and azimuthal velocity (v_θ). The geometrical shape of an epicyclic orbit is described by six constant parameters; semi-major axis (a), inclination (I_0), right ascension of ascending node (Ω_0), non singular parameters for undefined epicycle phase at perigee passage (needed for equatorial orbits) (ξ_p, η_p) and equator crossing time (t_E).

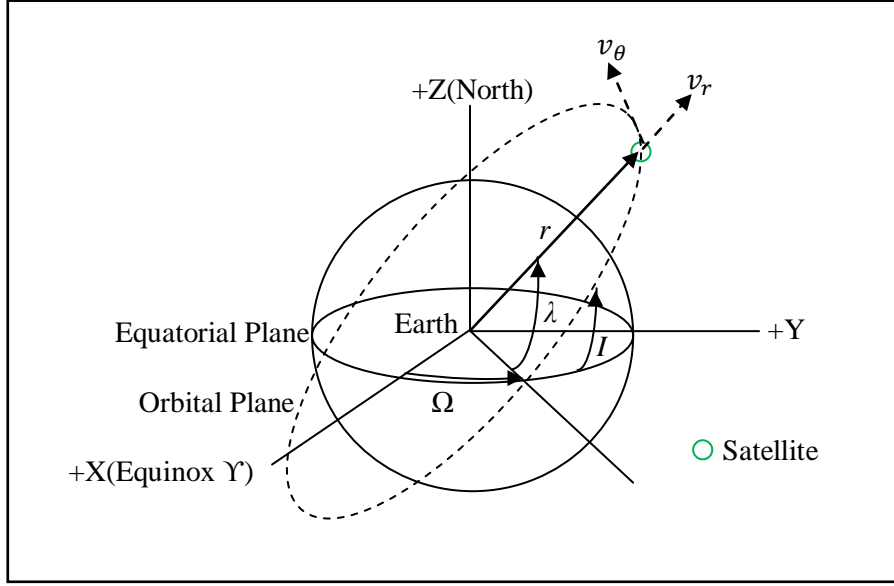


Figure 3-19: Geometrical representation of Epicycle coordinates of $(r, \lambda, I, \Omega, v_r, v_\theta)$ in ECI coordinate frame. Inclination (I) and right ascension of the ascending node (Ω) defines the orbital plane of a satellite and radial coordinate (r) and argument of latitude (λ) locate the position of the satellite on that plane. Radial velocity (v_r) and azimuthal velocity (v_θ) are also shown.

The mathematical expressions for the quantities are expressed in [2],[14]:

$$\begin{aligned}
 r &= a [1 + \rho - (\xi_p \cos \alpha + \eta_p \sin \alpha) + \Delta r_2 \cos 2\beta + \chi \sin \beta + \Delta_r] \\
 I &= I_0 + \Delta I_2 (1 - \cos 2\beta) + \Delta_I \\
 \Omega &= \Omega_0 + \vartheta \alpha + \Delta \Omega_2 \sin 2\beta + \Delta_\Omega \\
 \lambda &= \beta + 2[\xi_p \sin \alpha + \eta_p (1 - \cos \alpha)] - 2\chi (1 - \cos \beta) + \Delta \lambda_2 \sin 2\beta + \Delta_\lambda \\
 v_r &= an [(\xi_p \sin \alpha - \eta_p \cos \alpha) - (1 + \kappa)(2\Delta r_2 \sin 2\beta - \chi \cos \beta) + \dot{\Delta}_r] \\
 v_\theta &= rn [(1 + \kappa) + 2(\xi_p \cos \alpha + \eta_p \sin \alpha) + (1 + \kappa)(2\Delta \lambda_2 \cos 2\beta - 2\chi \sin \beta) + \dot{\Delta}_\lambda]
 \end{aligned} \tag{3.35}$$

where, r is the radius, I is the inclination, Ω represents the right ascension of the ascending node, λ is the argument of latitude, v_r is the radial velocity, v_θ is the azimuthal velocity, $\beta = (1 + \kappa)\alpha$, $\alpha = n(t - t_E)$, where t_E is the Equator passage time, Δx_2 is the short periodic coefficient due to J_2 , Δ_x is the short periodic coefficients due to higher zonal harmonic terms, long periodic variations in the orbit are described by χ . Other constants include, semi-major axis (a), inclination (I_0), right ascension of ascending node (Ω_0), (ξ_p, η_p) are the non singular parameters for the undefined epicycle phase at perigee passage where, $\xi_p = \frac{A}{a} \cos \alpha_p$, $\eta_p = \frac{A}{a} \sin \alpha_p$, $\alpha_p = n(t_p - t_E)$, t_p is the perigee passage time, A is the Epicycle amplitude and n denotes the mean motion. κ, ϑ are secular variations in the orbit. The

quantities ρ, κ, ϑ for J_2 are given by [2]:

$$\begin{aligned}\kappa_2 &= \frac{3}{4}J_2 \left(\frac{R_E}{a}\right)^2 (4 - 5 \sin^2 I_0) \\ \vartheta_2 &= -\frac{3}{2}J_2 \left(\frac{R_E}{a}\right)^2 \cos I_0 \\ \rho_2 &= -\frac{1}{4}J_2 \left(\frac{R_E}{a}\right)^2 (2 - 3 \sin^2 I_0)\end{aligned}\tag{3.36}$$

The short periodic coefficients for J_2 are provided in Appendix-C. In this analysis higher order zonal harmonic coefficients are not considered for secular and periodic variations in the orbit other than J_2 . Moreover, there are no long periodic effects in the orbit due to J_2 . Therefore one would neglect the χ term in Equation: 3.35. Thus, the differentials of epicycle coordinates (r, λ, I, Ω) from Equation: 3.35 after neglecting higher order zonal and χ terms are [14]:

$$\begin{aligned}\frac{\dot{r}}{an} &= (\xi_p \sin \alpha - \eta_p \cos \alpha) - (1 + \kappa)(2\Delta r_2 \sin 2\beta) \\ \frac{\dot{I}}{n} &= 2(1 + \kappa)\Delta I_2 \sin 2\beta \\ \frac{\dot{\Omega}}{n} &= \mathcal{G} + 2(1 + \kappa)\Delta \Omega_2 \cos 2\beta \\ \frac{\dot{\lambda}}{n} &= (1 + \kappa) + 2(\xi_p \cos \alpha + \eta_p \sin \alpha) + (1 + \kappa)(2\Delta \lambda_2 \cos 2\beta)\end{aligned}\tag{3.37}$$

By using epicyclic orbital coordinates of (r, I, Ω, λ) , the ECI position coordinates could be expressed as [14]:

$$\begin{aligned}X &= r(\cos \lambda \cos \Omega - \sin \lambda \cos I \sin \Omega) \\ Y &= r(\cos \lambda \sin \Omega + \sin \lambda \cos I \cos \Omega) \\ Z &= r \sin \lambda \sin I\end{aligned}\tag{3.38}$$

Using Equation: 3.35 and differentials of Equation: 3.37, the ECI velocity coordinates are [14]:

$$\begin{aligned}\dot{X} &= \dot{r}(c\lambda c\Omega - s\lambda cI s\Omega) + r\dot{I}s\lambda sI s\Omega - r\dot{\Omega}(c\lambda s\Omega + s\lambda cI c\Omega) - r\dot{\lambda}(s\lambda c\Omega + c\lambda cI s\Omega) \\ \dot{Y} &= \dot{r}(c\lambda s\Omega + s\lambda cI c\Omega) - r\dot{I}s\lambda sI c\Omega + r\dot{\Omega}(c\lambda c\Omega - s\lambda cI s\Omega) - r\dot{\lambda}(s\lambda s\Omega - c\lambda cI c\Omega) \\ \dot{Z} &= \dot{r}s\lambda sI + r\dot{I}s\lambda cI + r\dot{\lambda}c\lambda sI\end{aligned}\tag{3.39}$$

where, “c” and “s” stands for sine and cosine functions. In this model the full orbital evolution equations (Equation: 3.35) determine the motion once we know the six epicycle parameters.

Therefore, one would consider epicycle orbital parameters \mathbf{x}_0 to be estimated, expressed as:

$$\mathbf{x}_0 = (a, \xi_P, \eta_P, I_0, \Omega_0, \alpha_0)^T \quad (3.40)$$

where, α_0 is the initial epicycle phase which is analogous to M_0 is the mean anomaly at epoch in classical orbital elements. The time history of ECI position and velocity for numerical trajectory are obtained by integrating equations of motion (Equation: 3.32) as previously done for analysis of Kepler's equation. The partial derivative matrix for this estimation problem (see Appendix-D for components) is expressed as [14]:

$$\mathbf{F} = \frac{\partial \mathcal{P}(\mathbf{x}_0, t)}{\partial \mathbf{x}_0} = \frac{\partial \mathcal{P}}{\partial \mathbf{y}(t)} \frac{\partial \mathbf{y}(t)}{\partial \mathbf{x}_0} \quad (3.41)$$

where, $\frac{\partial \mathcal{P}(\mathbf{x}_0, t)}{\partial \mathbf{y}(t)}$ is the partial derivative matrix of epicycle trajectory (in ECI coordinates) with respect to epicycle coordinates $\mathbf{y}(t)$, $\frac{\partial \mathbf{y}(t)}{\partial \mathbf{x}_0}$ is the State Transition Matrix (STM) / partial derivative matrix of epicycle coordinates with respect to epicycle parameters and vector " $\mathbf{y}(t)$ " for the problem, consisting of epicycle coordinates is expressed as:

$$\mathbf{y} = (r, I, \Omega, \lambda, \dot{r}, \dot{I}, \dot{\Omega}, \dot{\lambda})^T \quad (3.42)$$

The initial conditions of the reference orbit are same as expressed in Equation: 3.33. Epicycle orbital parameters (Equation: 3.40) were found out by using the estimator for orbital data generated over time span of 10 orbital periods. The optimal choice of these parameters is shown in Table: 3-5.

Orbital Elements	IC of Numerical Trajectory \mathbf{x}_0	IC of Analytical Trajectory (Output of estimator - $\hat{\mathbf{x}}_0$)
a	6863.100 km	6863.10048 km
ξ_P	0.0001	1.0021×10^{-4}
η_P	0	3.2049×10^{-10}
I_0	98 deg	97.9999 deg
Ω_0	0 deg	5.932×10^{-7} deg
α_0	0 deg	4.0882×10^{-8} deg

Table 3-5: Comparison of IC for numerical and analytical trajectory for J_2 perturbed 2 body equation once compared with analytical epicycle orbit.

With IC estimates $\hat{\mathbf{x}}_0$ presented in Table: 3-5 the epicycle orbit of a satellite is propagated forward using Equations: 3.38 and 3.39 and then converted into LVLH frame of satellite propagated through numerical integration of Equation: 3.32 using RK-4. The results are presented in Figure: 3-20 and 3-21.

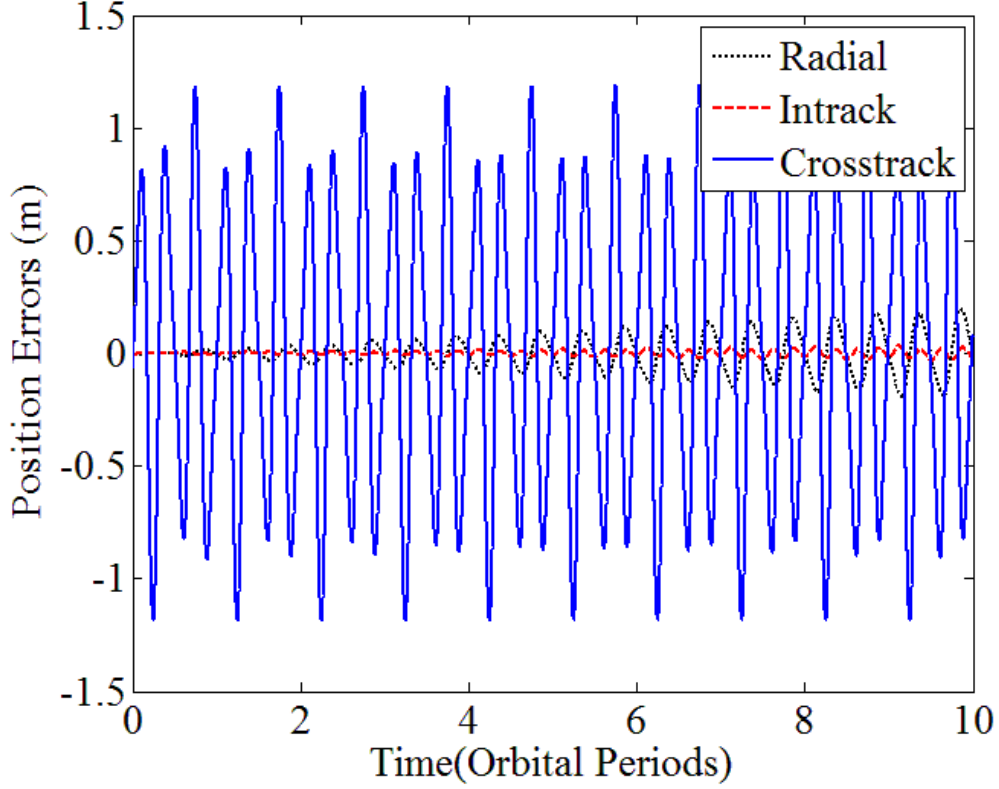


Figure 3-20: Time history of position errors for epicycle orbit compared with numerical trajectory of J_2 perturbed 2 *body* equation.

The errors in both positions and velocity are considerably small and show no divergence over time. Table: 3-6 summarizes the maximum errors for position and velocity the LVLH coordinate frame. The maximum in-track and radial positional errors are about 0.034 m and 0.19 m respectively over 10 orbital periods which is a significant improvement over the solution of Kepler's Equation (see Table: 3-4). However, the maximum cross track error is about 1.19 meters which is due to the difference in periodic variations of RAAN and the inclination. The error in velocity plots is also low on the order of 0.0001-0.0043 m/s. The epicycle propagation Equations: 3.35 and 3.37 also take into account the second order epicycle coefficients for J_2 i.e., J_2^2 [2] (see Appendix-C for details). Thus, the Epicycle

model is quite accurate and shows improvement in fidelity with a proper choice of initial conditions.

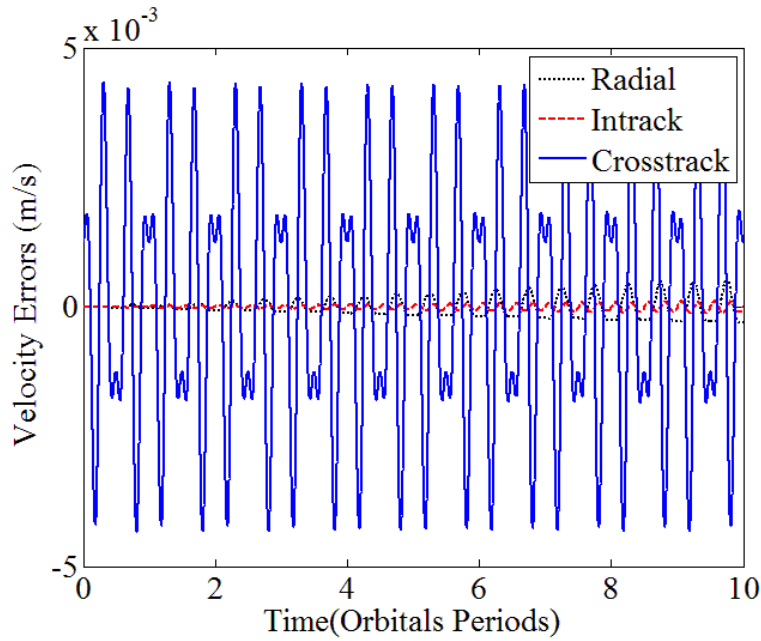


Figure 3-21: Time history of velocity errors for epicycle orbit compared with numerical trajectory of J_2 perturbed 2 body equation.

The position and velocity errors for the Epicycle model without modifying the parameters are now calculated to compare the effectiveness of choosing appropriate parameters for orbital propagations. The results are shown in Figures: 3-22 and 3-23. The positional errors for in-track direction show a secular drift and an increased growth of periodic errors in cross track directions. The in-track errors are due to inappropriate choice of the semi-major axis “ a ” and “ α_0 ” and cross track errors due to different I and Ω . Therefore the appropriate choice of parameters is crucial.

EP used for propagation of epicycle orbit	Position (m)			Velocity (m/s)		
	R	I	C	R	I	C
\mathbf{x}_0	2	47.86	2.14	0.0023	.0042	0.0037
$\hat{\mathbf{x}}_0$	0.19	0.034	1.19	0.0005	.0001	0.0043
EP – Epicycle Parameters R – Radial, I – In-track, C – Cross-track \mathbf{x}_0 – EP for Numerical trajectory $\hat{\mathbf{x}}_0$ – Estimated EP for Analytic trajectory						

Table 3-6: Summary of the Maximum Absolute Position and Velocity Errors in LVLH Coordinates over 10 Orbital Periods for the Epicycle Model.

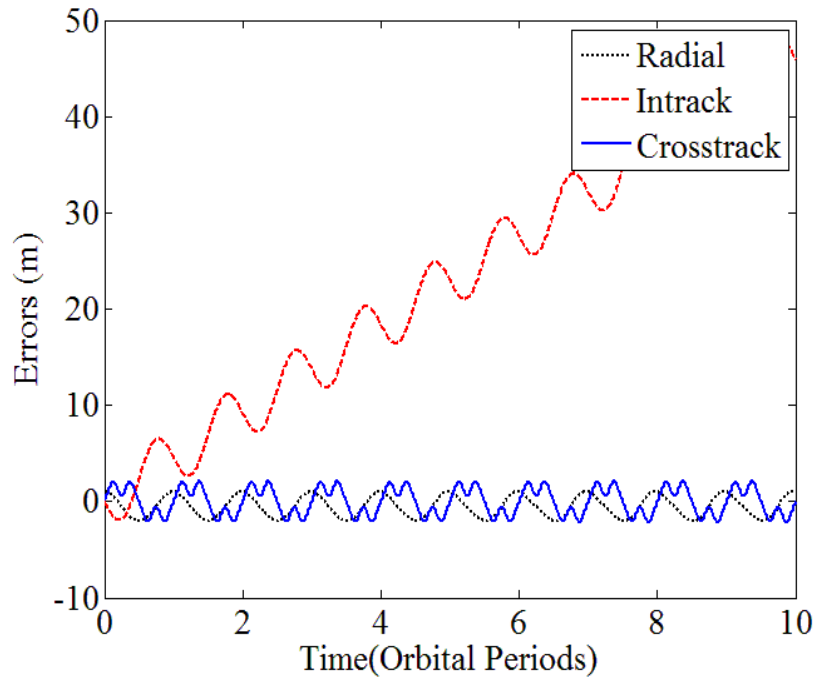


Figure 3-22: Time history of position errors for epicycle orbit compared with numerical trajectory of J_2 perturbed 2 *body* equation without estimating IC.

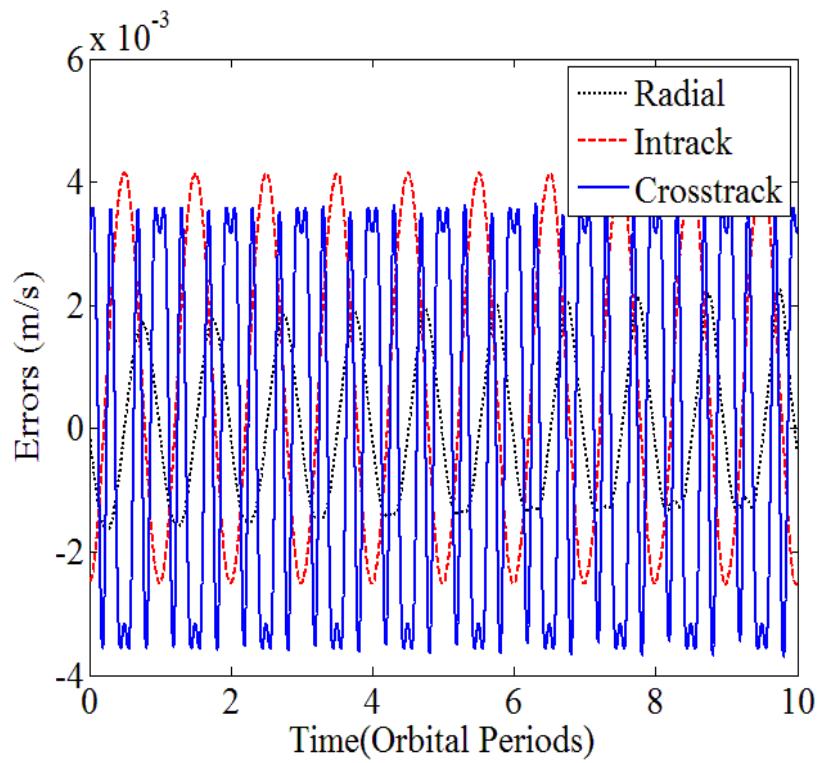


Figure 3-23: Time history of velocity errors for epicycle orbit compared with numerical trajectory of J_2 perturbed 2 *body* equation without estimating IC.

3.5.3 Conclusion

In this section the analytical solution for 2 *body* equations of motion due to spherical and non-spherical geopotential has been examined. Solutions of the equation of Kepler provide the fundamental concept of orbital motion. The solution is almost exact for equations of motion without perturbations due to non-spherical Earth. However, for non-spherical Earth, the solution of Kepler's equation is insufficient to capture the true orbital dynamics characterized by secular and periodic variations which is the main source of its process noise. The analytic solution expressed as epicyclic motion of a satellite around an oblate Earth by Hashida and Palmer captures the secular and periodic variations in the orbit of satellite sufficiently well. Improvement in efficiency due to correct choice of parameters $\hat{\mathbf{x}}_0$ for this orbit has also been demonstrated. By using appropriate IC one is able to reduce positional errors considerably in all the three directions. For Kepler's equation the error is reduced by 60% in radial, 99.7% in in-track, 49.9% in cross track respectively. For Epicycle model it is reduced by 90.5% in radial, 99.92% in in-track, and 44.39% in cross track respectively.

3.6 Relative Motion between Satellites

Recent interest in formation of satellites, in wide range of space missions [17],[86] has revived the interest in development and use of relative motion models [18],[19]. In this section a relative motion model of two satellites will be described. Briefly, described in Section: 2.4, the basic relative orbital motion is defined for a formation of two satellites where the motion of one of the satellite known as deputy is considered with respect to another known as chief satellite [88]. There are different choices of relative motion coordinate systems and reference frames for this description. A geocentric ECI and the chief centred Local LVLH coordinate system are two choices shown in Figure: 3-24. The LVLH coordinate system is fixed to the chief satellite and the relative motion of a deputy satellite is described in three directions i.e., the motion along x , y , and z is referred as radial, In-track and cross-track motion, respectively. LVLH coordinate system is a good choice for visualizing the relative orbits. In ECI coordinate system the relative motion can be obtained by integrating the two sets of Equation: 3.32, one for chief and one for deputy. The inertial relative displacement and velocity vectors are expressed as [88]:

$$\begin{aligned}\delta\vec{r} &= \vec{r}_d - \vec{r}_c \\ \delta\vec{v} &= \vec{v}_d - \vec{v}_c\end{aligned}\tag{3.43}$$

where, the subscript “c” and “d” denote chief and deputy satellite respectively.

The relative motion between the two satellites can be transformed into LVLH from ECI coordinates using a transformation matrix $\mathfrak{R}^{L/E}$ defined as follows [12]:

$$\mathfrak{R}^{L/E} = \begin{bmatrix} \hat{e}_R^T \\ \hat{e}_I^T \\ \hat{e}_C^T \end{bmatrix} \quad (3.44)$$

where, R stands for radial, I denotes in-track, C is the cross-track, $\mathfrak{R}^{L/E}$ is the transformation matrix which transforms ECI (E) coordinates into LVLH (L) coordinates and the unit vectors $\hat{e}_{(.)}$ are defined as:

$$\begin{aligned} \hat{e}_R &= \frac{\vec{r}_c}{r_c} \\ \hat{e}_I &= \hat{e}_C \times \hat{e}_R \\ \hat{e}_C &= \frac{\vec{r}_c \times \vec{v}_c}{|\vec{r}_c \times \vec{v}_c|} \end{aligned} \quad (3.45)$$

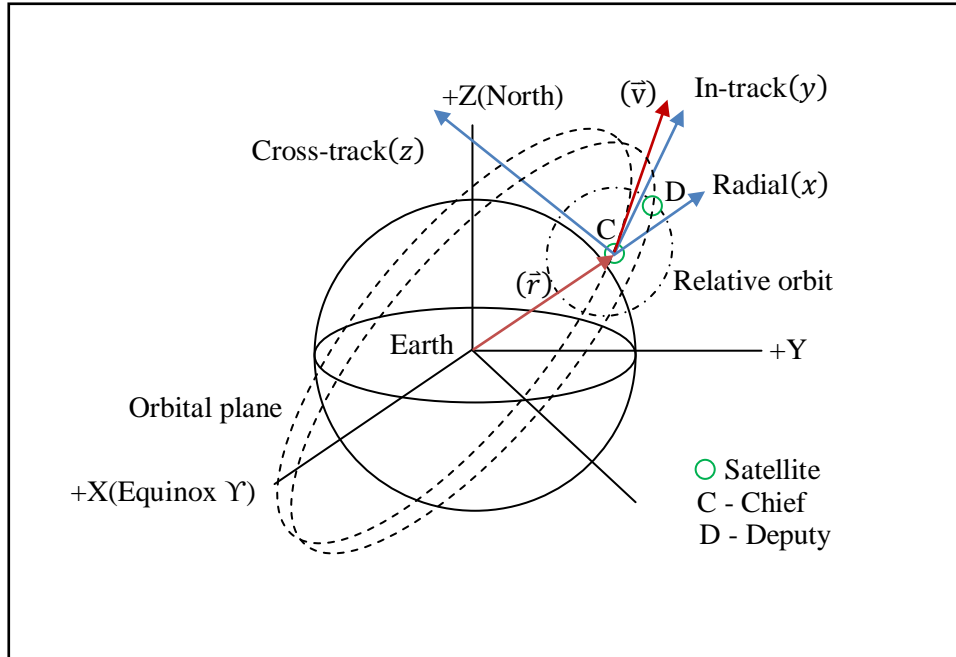


Figure 3-24: Illustration of the satellite relative motion coordinate system, Local Vertical Local Horizontal (LVLH) with reference to Chief (C) satellite. (x) axis points from the Earth's centre along the radius vector towards the satellite as it moves along the orbit. (y) axis points in the direction of velocity vector (not necessarily parallel) and is perpendicular to radius vector. The (z) axis is normal to the orbital plane. Relative motion of Deputy (D) satellite can be expressed in Chief (C) satellite centred LVLH reference frame.

By using Equation: 3.43-44, the relative position in LVLH coordinates is obtained as:

$$\vec{r}_{LVLH} = \mathfrak{R}^{L/E} \delta \vec{r} \quad (3.46)$$

The relative velocity in LVLH coordinates is expressed using the principal of kinematics and Equation: 3.43 and 3.46 as:

$$\vec{v}_{LVLH} = \delta \vec{v} - \vec{\omega}_c \times \vec{r}_{LVLH} \quad (3.47)$$

where,

$$\vec{\omega}_c = \frac{\vec{r}_c \times \vec{v}_c}{r_c^2} \quad (3.48)$$

3.7 Analysis of Relative Motion

In this section analysis of two linearized satellite relative motion models will be carried out. Firstly, a relative motion model described for chief and deputy satellite orbits with assumptions of spherically symmetric geopotential, circular orbit of chief satellite and linearized differential gravity acceleration will be analyzed. The model is termed as HCW Equations [18],[19]. Secondly, analysis of a relative motion model for satellites under non-spherical geopotential for zonal harmonic terms up to J_2 will be undertaken. This model is termed as J_2 modified HCW by Shweighthart and Sedwick (SS) [20]. The nonlinear satellite relative motion model developed in Equations: 3.46 and 3.47 is being considered as reference (true) relative motion. The orbit of chief satellite is chosen as Synthetic Aperture RADAR (SAR) Lupe-1 sun-synchronous orbit. The initial conditions for this satellite are obtained from the North American Air Defence Command (NORAD) Two Line Element (TLE) set expressed as; $a = 6863.100$ km, $e = 0.0015961$, $I = 98.1794$ deg, $\Omega = 84.4914$ deg, $\omega = 2.2798$ deg, $M_0 = 133.5407$ deg [101]. NORAD maintains the TLE set for each operational satellite and for the large non-operational satellite / debris orbiting Earth (for details see Ref. [101]). The deputy Satellite is selected to be in free orbit ellipse relative orbit (natural closed orbital path of satellite in a formation) with relative orbit amplitude $A = 50$ m [20] as shown in Figure: 3-25.

The fundamental idea in this analysis is the optimal selection of IC for relative orbit described by linearized equations of motion. These IC would minimize the difference between the reference (true) and linearized relative motion. The selection of IC using GLDC scheme provides an optimal choice for such condition. However, the relative orbital models due to difference in assumptions on true

nonlinear dynamics of chief satellite would present a varying fidelity when compared with truth orbit.

3.7.1 Hill Clohessy Wiltshire Model

The HCW model for satellite relative motion are set of three second order linear differential equations expressed as [19],[18],[12]:

$$\begin{aligned}\ddot{x} - 2n\dot{y} - 3n^2x &= 0 \\ \ddot{y} + 2n\dot{x} &= 0 \\ \ddot{z} + n^2z &= 0\end{aligned}\tag{3.49}$$

where, x , y and z are relative motion coordinates in LVLH frame of reference, centred on chief satellite.

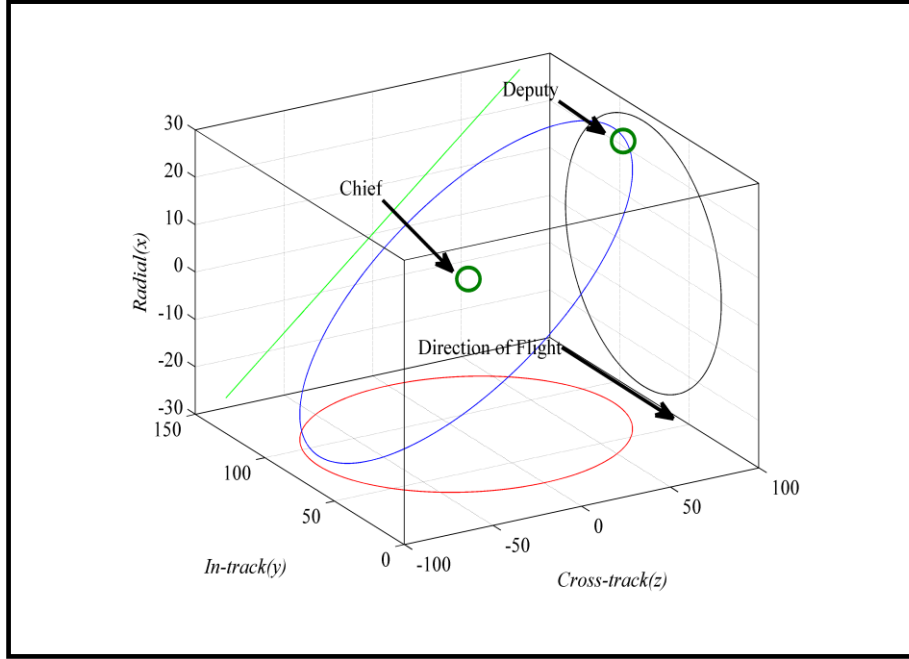


Figure 3-25: Geometry of the free orbit ellipse for relative motion of chief and deputy satellites (drawn in blue colour). The projection of the deputy satellite orbit (drawn in blue colour) on y - z plane forms a circle (drawn in red colour), projection on x - y plane forms (2×1) ellipse (drawn in black colour) and projection on x - z plane forms a line (drawn in green colour).

The analytical solutions to HCW equations admit bounded periodic orbits (subject to suitable initial conditions) and are given as [12]:

$$\begin{aligned}
 x(t) &= \frac{\dot{x}_0}{n} \sin(nt) - (3x_0 + \frac{2\dot{y}_0}{n}) \cos(nt) + (4x_0 + \frac{2\dot{y}_0}{n}) \\
 y(t) &= (6x_0 + \frac{4\dot{y}_0}{n}) \sin(nt) + \frac{2\dot{x}_0}{n} \cos(nt) - (6nx_0 + 3\dot{y}_0)t + (y_0 - \frac{2\dot{x}_0}{n}) \\
 z(t) &= z_0 \cos(nt) + \frac{\dot{z}_0}{n} \sin(nt) \\
 \dot{x}(t) &= \dot{x}_0 \cos(nt) + (3nx_0 + 2\dot{y}_0) \sin(nt) \\
 \dot{y}(t) &= (6nx_0 + 4\dot{y}_0) \cos(nt) - 2\dot{x}_0 \sin(nt) - (6nx_0 + 3\dot{y}_0) \\
 \dot{z}(t) &= -z_0 n \sin(nt) + \dot{z}_0 \cos(nt)
 \end{aligned} \tag{3.50}$$

where, $\mathbf{x}_0 = [x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, \dot{z}_0]^T$ is the IC for the HCW relative orbit. Given any IC the relative motion coordinates of a deputy satellite can be obtained at any time “ t ”:

$$\begin{aligned}
 \vec{r}_{LVLH}^p(t) &= [x(t) \ y(t) \ z(t)]^T \\
 \vec{v}_{LVLH}^p(t) &= [\dot{x}(t) \ \dot{y}(t) \ \dot{z}(t)]^T
 \end{aligned} \tag{3.51}$$

where, superscript “p” in Equation: 3.51 denotes the analytical solution for the relative motion of a satellite. One may notice the secular drift term in the expression for in-track (y) solution in Equation: 3.50. In order to obtain the zero secular drift term, the initial condition for \dot{y}_0 would be obtained as:

$$\dot{y}_0 = -2nx_0 \tag{3.52}$$

Since this analysis is based on equations of motion given in ECI coordinate frame therefore; the relative motion of a satellite (given in Equation 3.50) is converted into ECI coordinates. Essentially, a reverse procedure from Equation: 3.46-3.47 is adopted to acquire these coordinates:

$$\begin{aligned}
 \vec{r}_d^p &= (\mathfrak{R}^{L/E})^{-1} \vec{r}_{LVLH}^p + \vec{r}_c \\
 \vec{v}_d^p &= \vec{v}_{LVLH}^p + \vec{\omega}_c \times \vec{r}_{LVLH}^p + \vec{v}_c
 \end{aligned} \tag{3.53}$$

The components $\vec{r}_d^p = [X_d, Y_d, Z_d]^T$ for ECI position coordinates can be expressed as:

$$\begin{aligned}
X_d &= \hat{e}_{RX} ((4 - 3 \cos nt)x_0 + \dot{x}_0 / n(\sin nt) + \frac{2(1 - \cos nt)}{n} \dot{y}_0) + \hat{e}_{IX} (-6(nt - \sin nt)x_0 \\
&- 2(1 - \cos nt) \frac{\dot{x}_0}{n} + y_0 - \frac{(3nt - 4 \sin nt)}{n} \dot{y}_0) + \hat{e}_{CX} (z_0 \cos nt + \dot{z}_0 / n(\sin nt)) + X_c \\
Y_d &= \hat{e}_{RY} ((4 - 3 \cos nt)x_0 + \dot{x}_0 / n(\sin nt) + \frac{2(1 - \cos nt)}{n} \dot{y}_0) + \hat{e}_{IY} (-6(nt - \sin nt)x_0 \\
&- 2(1 - \cos nt) \frac{\dot{x}_0}{n} + y_0 - \frac{(3nt - 4 \sin nt)}{n} \dot{y}_0) + \hat{e}_{CY} (z_0 \cos nt + \dot{z}_0 / n(\sin nt)) + Y_c \\
Z_d &= \hat{e}_{RZ} ((4 - 3 \cos nt)x_0 + \dot{x}_0 / n(\sin nt) + \frac{2(1 - \cos nt)}{n} \dot{y}_0) + \hat{e}_{IZ} (-6(nt - \sin nt)x_0 \\
&- 2(1 - \cos nt) \frac{\dot{x}_0}{n} + y_0 - \frac{(3nt - 4 \sin nt)}{n} \dot{y}_0) + \hat{e}_{CZ} (z_0 \cos nt + \dot{z}_0 / n(\sin nt)) + Z_c
\end{aligned} \tag{3.54}$$

where, the unit vectors \hat{e}_R , \hat{e}_I and \hat{e}_C are obtained from Equation: 3.45 and $[X_c \ Y_c \ Z_c]^T$ are ECI position coordinates of chief. The ECI velocity coordinates $\vec{v}_d^p = [\dot{X}_d, \dot{Y}_d, \dot{Z}_d]^T$ are expressed as:

$$\begin{aligned}
\dot{X}_d &= 3x_0 n \sin nt + \dot{x}_0 \cos nt + 2 \sin nt \dot{y}_0 + \{\mathbf{A}_2(z_0 \cos nt + \frac{\dot{z}_0}{n} \sin nt) - \\
&\mathbf{A}_3(-6(nt - \sin nt)x_0 - 2(1 - \cos nt) \frac{\dot{x}_0}{n} + y_0 - (3nt - 4 \sin nt)/n(\dot{y}_0)) + \dot{X}_c \\
\dot{Y}_d &= -6(1 - \cos nt)nx_0 - 2 \sin nt \dot{x}_0 - (3 - 4 \cos nt)\dot{y}_0 - \{\mathbf{A}_1(z_0 \cos nt + \frac{\dot{z}_0}{n} \sin nt) \\
&- \mathbf{A}_3((4 - 3 \cos nt)x_0 + \frac{\dot{x}_0}{n} \sin nt + 2(1 - \cos nt)/n(\dot{y}_0))\} + \dot{Y}_c \\
\dot{Z}_d &= -\sin(nt)nz_0 + \dot{z}_0 \cos nt + \{\mathbf{A}_1(-6(nt - \sin nt)x_0 - 2(1 - \cos nt) \frac{\dot{x}_0}{n} + y_0 \\
&- \frac{(3nt - 4 \sin nt)}{n} \dot{y}_0) - \mathbf{A}_2(4 - 3 \cos nt)x_0 + \frac{\dot{x}_0}{n} \sin nt + \frac{2(1 - \cos nt)}{n} \dot{y}_0) + \dot{Z}_c
\end{aligned} \tag{3.55}$$

where, \mathbf{A}_1 , \mathbf{A}_2 , and \mathbf{A}_3 are components of the angular velocity vector of the chief satellite $\vec{\omega}_c$ is expressed in Equation: 3.48 and $[\dot{X}_c \ \dot{Y}_c \ \dot{Z}_c]^T$ are ECI velocity coordinates of chief satellite. Equations: 3.54-3.55 would be considered as an analytical description of the deputy satellite. The partial derivative matrix \mathbf{F} for the estimation of IC for a deputy satellite is obtained as:

$$\mathbf{F} = \frac{\partial \mathcal{P}(\mathbf{x}_0, t)}{\partial \mathbf{x}_0} \tag{3.56}$$

The partials for ECI position coordinates are expressed as:

$$\frac{\partial X_d}{\partial \mathbf{x}_0} = \begin{bmatrix} \hat{e}_{RX}(4 - 3 \cos nt)x_0 + \hat{e}_{IX}(-6(nt - \sin nt)) \\ \hat{e}_{IX} \\ \hat{e}_{CX}(\cos nt) \\ \hat{e}_{RX} \frac{\sin nt}{n} + \hat{e}_{IX} \left(\frac{-2(1 - \cos nt)}{n} \right) \\ e_{RX} \left(\frac{2(1 - \cos nt)}{n} \right) + e_{IX} \frac{-(3nt - 4 \sin nt)}{n} \\ e_{CX} \left(\frac{\sin nt}{n} \right) \end{bmatrix}^T \quad (3.57)$$

$$\frac{\partial Y_d}{\partial \mathbf{x}_0} = \begin{bmatrix} \hat{e}_{RY}(4 - 3 \cos nt) + \hat{e}_{IY}(-6(nt - \sin nt)) \\ \hat{e}_{IY} \\ \hat{e}_{CY} \cos nt \\ \hat{e}_{RY} \frac{\sin nt}{n} + \hat{e}_{IY} \left(\frac{-2(1 - \cos nt)}{n} \right) \\ e_{RY} \left(\frac{2(1 - \cos nt)}{n} \right) + e_{IY} \frac{-(3nt - 4 \sin nt)}{n} \\ e_{CY} \left(\frac{\sin nt}{n} \right) \end{bmatrix}^T$$

$$\frac{\partial Z_d}{\partial \mathbf{x}_0} = \begin{bmatrix} \hat{e}_{RZ}(4 - 3 \cos nt) + \hat{e}_{IZ}(-6(nt - \sin nt)) \\ \hat{e}_{IZ} \\ \hat{e}_{CZ} \cos nt \\ \hat{e}_{RZ} \frac{\sin nt}{n} + \hat{e}_{IZ} \left(\frac{-2(1 - \cos nt)}{n} \right) \\ e_{RZ} \left(\frac{2(1 - \cos nt)}{n} \right) + e_{IZ} \frac{-(3nt - 4 \sin nt)}{n} \\ e_{CZ} \left(\frac{\sin nt}{n} \right) \end{bmatrix}^T$$

The partials for ECI velocity coordinates are expressed as:

$$\begin{aligned} \frac{\partial \dot{X}_d}{\partial \mathbf{x}_0} &= \begin{bmatrix} 3n \sin nt - \mathbf{A}_3(-6(nt - \sin nt)) \\ -\mathbf{A}_3 \\ \mathbf{A}_2 \cos nt \\ \cos nt - \mathbf{A}_3 \left(\frac{-2(1 - \cos nt)}{n} \right) \\ 2 \sin nt - \mathbf{A}_3 \frac{-(3nt - 4 \sin nt)}{n} \\ \mathbf{A}_2 \frac{\sin nt}{n} \end{bmatrix}^T \quad (3.58) \\ \\ \frac{\partial \dot{Y}_d}{\partial \mathbf{x}_0} &= \begin{bmatrix} -6(1 - \cos nt)n + \mathbf{A}_3(4 - 3 \cos nt) \\ 0 \\ -\mathbf{A}_1 \cos nt \\ -2 \sin nt + \mathbf{A}_3 \frac{\sin nt}{n} \\ -(3 - 4 \cos nt) + \mathbf{A}_3 \left(\frac{-2(1 - \cos nt)}{n} \right) \\ -\mathbf{A}_1 \frac{\sin nt}{n} \end{bmatrix}^T \\ \\ \frac{\partial \dot{Z}_d}{\partial \mathbf{x}_0} &= \begin{bmatrix} \mathbf{A}_1(-6(nt - \sin nt)) - \mathbf{A}_2(4 - 3 \cos nt) \\ \mathbf{A}_1 \\ -n \sin nt \\ \mathbf{A}_1 \left(\frac{-2(1 - \cos nt)}{n} \right) - \mathbf{A}_2 \frac{\sin nt}{n} \\ \mathbf{A}_1 \frac{-(3nt - 4 \sin nt)}{n} - \mathbf{A}_2 \left(\frac{-2(1 - \cos nt)}{n} \right) \\ \cos nt \end{bmatrix}^T \end{aligned}$$

where, \mathbf{x}_0 is the state vector to be estimated and is given as $\mathbf{x}_0 = [x_0, y_0, z_0, \dot{x}_0, \dot{y}_0, \dot{z}_0]^T$. Similarly the analysis for absolute satellite orbital dynamics, one would now compare the HCW relative orbital model with the reference nonlinear relative motion developed in Equation: 3.46-3.47. In view of the estimation scheme developed in Section: 3.2, one would require orbital data for reference and

analytical trajectory of a deputy satellite. The data of the reference deputy satellite for 10 orbital periods was generated by numerically integrating Equation: 3.32 using RK-4 with a step size of 5 sec. The orbital data for analytical trajectory of the same deputy satellite was obtained using analytic solutions (Equation: 3.54-3.55). The initial conditions of a free orbit ellipse periodic orbit as shown in Figure: 3-25 were selected for the deputy satellite with initial relative orbit amplitude $A = 50$ m and initial phase (ϕ) = 56 deg (in this section reuse of " ϕ " for initial phase instead of geocentric latitude is done) as shown in Figure: 3-26.

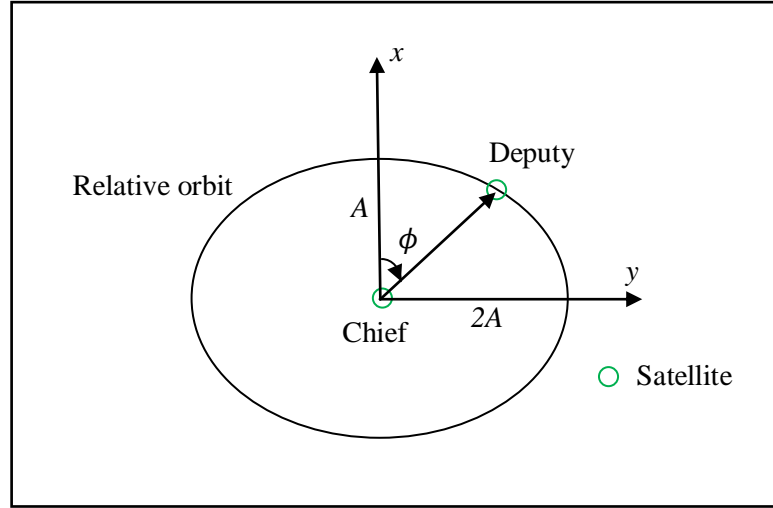


Figure 3-26: Illustration of “free orbit ellipse” relative orbit in x - y plane forming (2×1) ellipse with amplitude $A = 50$ m and initial phase $\phi = 56$ deg (drawn not to scale).

The LVLH coordinates of the deputy satellite are converted into ECI Position and Velocity using Equation 3.53. This would be used as initial conditions for generating numerical trajectory of the deputy satellite. Using the estimation algorithm the optimal IC $\hat{\mathbf{x}}_0$ for orbit of deputy is found out (see Table: 3-7).

IC in LVLH Coordinates frame	IC of Numerical Trajectory \mathbf{x}_0	IC of Analytical Trajectory (Output of estimator - $\hat{\mathbf{x}}_0$)
x_0	$A \cos \phi = 27.95$ m	33.22 m
y_0	$2A \sin \phi = 82.90$ m	87.18 m
z_0	$2A \sin \phi = 82.90$ m	83.57 m
\dot{x}_0	0.0 m/s	0.2722 m/s
\dot{y}_0	-0.0620 m/s	-0.00241 m/s
\dot{z}_0	0.0 m/s	0.0066 m/s

Table 3-7: Comparison of IC for numerical and analytical trajectory for HCW equations compared with J_2 perturbed full nonlinear relative motion equations.

The estimated IC expressed in Table: 3-7 are then used for propagation of the orbit of the deputy satellite using Equation: 3.50. The reference relative motion of the deputy satellite obtained from Equation: 3.46-3.47 is considered for comparison. In other words, the errors in ECI coordinates between the two orbital descriptions (analytical and numerical) for a deputy satellite are converted into LVLH frame of deputy satellite whose orbital data is obtained from numerical integration of equations of motion which could now be considered as chief satellite. These results are shown in Figure: 3-27 and 3-28.

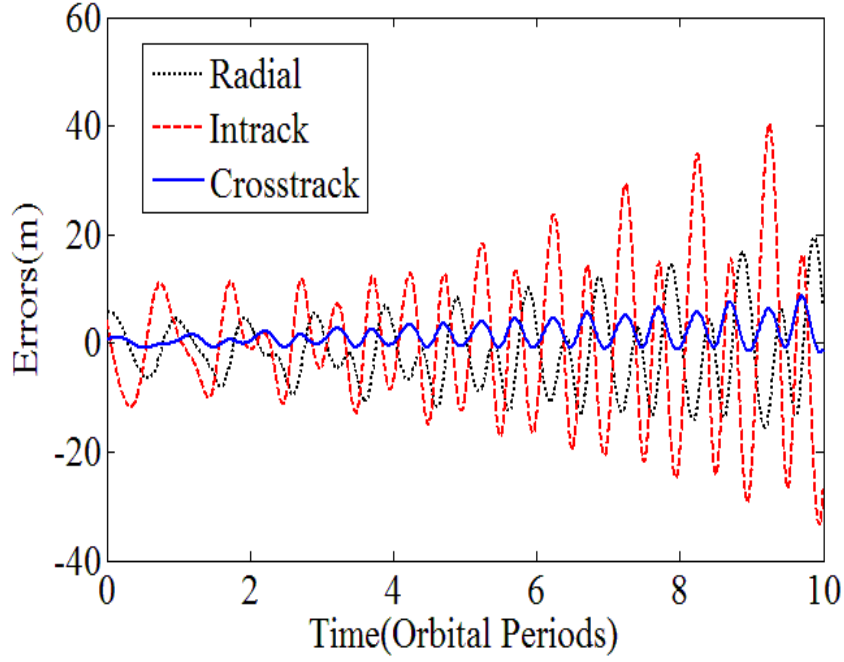


Figure 3-27: Time history of position errors for HCW equations using optimal initial conditions $\hat{\mathbf{x}}_0$ in LVLH coordinate frame.

The error plots (Figure: 3-27 and 3-28) indicate growth of errors in all three directions. Table: 3-8 summarizes errors in position and velocity coordinates. The worst case error is observed in in-track direction. The error is periodically increasing with a secular drift. The error has gone up to 40 m in 10 orbital periods, owing to inability of HCW equations to capture the difference in the orbital energies of satellite experiencing J_2 which is due to the difference in the semi-major axis “ a ”. Bearing in mind the precession of the orbit of satellite experiencing J_2 around the North Pole of Earth and a continuous nodal drift, cross track motion is visualized. As stated earlier that the cross track motion is solely dependent on the difference in the inclination and nodal separation of the two orbital planes which does not remain constant under the influence of J_2 . Thus there is an increase in the error in the cross-

track amplitude of maximum of 8 m in 10 orbital periods and continuous drift as viewed in the simulation results. The radial direction errors are also periodically increasing as the instantaneous semi-major axis of the perturbed orbit is also varying which is not captured by HCW equations. - However, the errors are about maximum of 19 m in 10 orbital periods. The velocity plots indicate rise in error periodically over the experimental time span. This indicates the HCW equations are not a true representative of nonlinear relative velocities especially with J_2 . The maximum error of 0.06 m/sec is observed in in-track direction. The velocity errors of radial and cross track are periodically increasing with maximum absolute error of 0.02 m/sec.

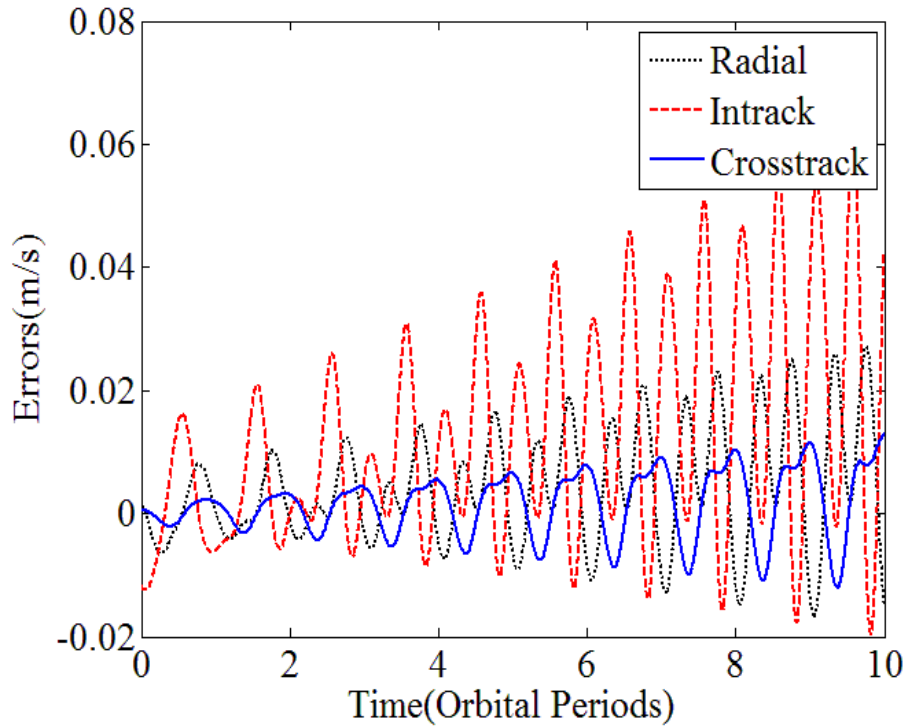


Figure 3-28: Time history of velocity errors HCW equations using optimal initial conditions $\hat{\mathbf{x}}_0$ in LVLH coordinate frame.

The errors in the analytic and true relative motion without using estimated initial conditions are shown in Figure: 3-29 and 3-30. The position errors clearly indicate breakdown of HCW solutions when compared with the true nonlinear relative motion. Moreover, the sensitivity of these solutions to IC is now clearly obvious. Errors in km are observed in in-track direction owing to differences in orbital energies of satellites perturbed by J_2 . The simplicity of HCW equations makes it the most favourable choice for the relative motion analysis. The analysis under different choices of the chief orbit is now being looked into. The most important orbital parameters are semi-major axis “ a ”, inclination “ I_0 ” and eccentricity “ e ”.

IC used for propagation of HCW Equations	Position (m)			Velocity (m/s)		
	R	I	C	R	I	C
\mathbf{x}_0	406	12663	14	0.32	0.82	.01
$\hat{\mathbf{x}}_0$	19	40	8	0.02	0.06	0.01
R – Radial, I – In-track, C – Cross-track						
\mathbf{x}_0 – IC for Numerical Trajectory						
$\hat{\mathbf{x}}_0$ – Estimated IC (Output of Estimator)						

Table 3-8: Summary of Maximum Absolute Position and Velocity Errors in LVLH Coordinates over 10 Orbital Periods for HCW Model

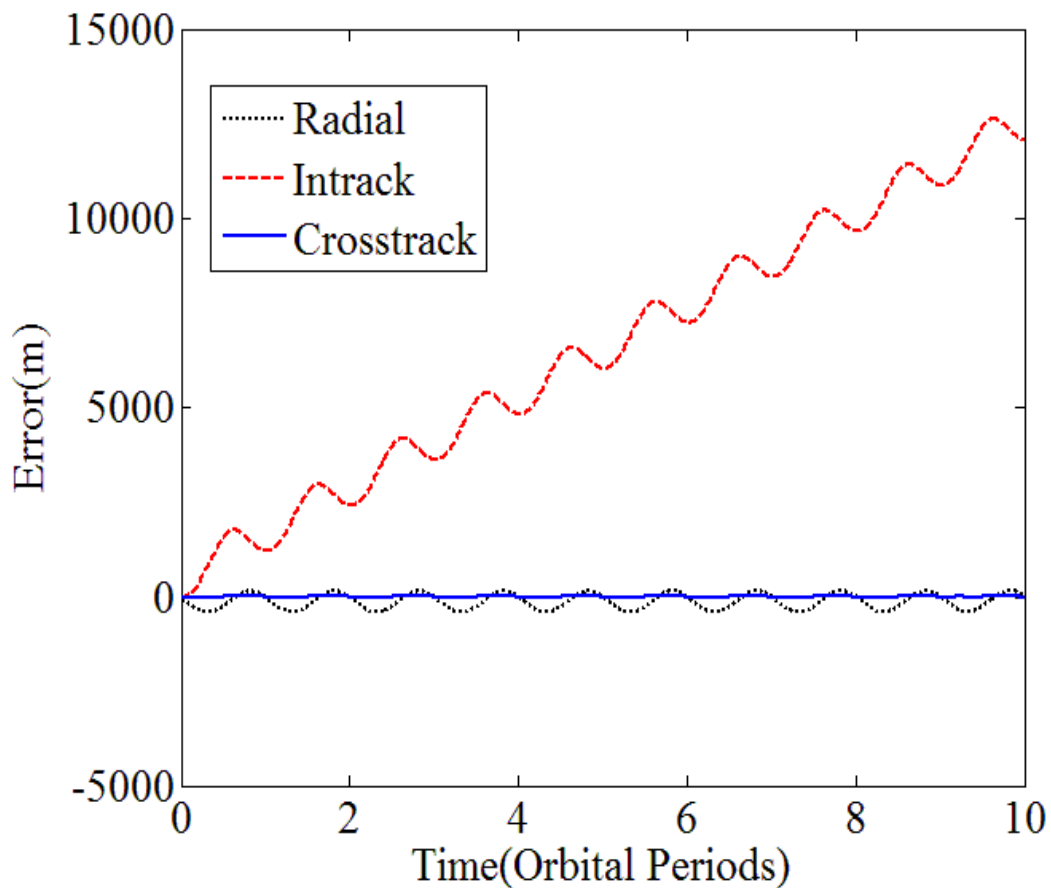


Figure 3-29: Time history of position errors HCW equations without using estimated initial conditions.

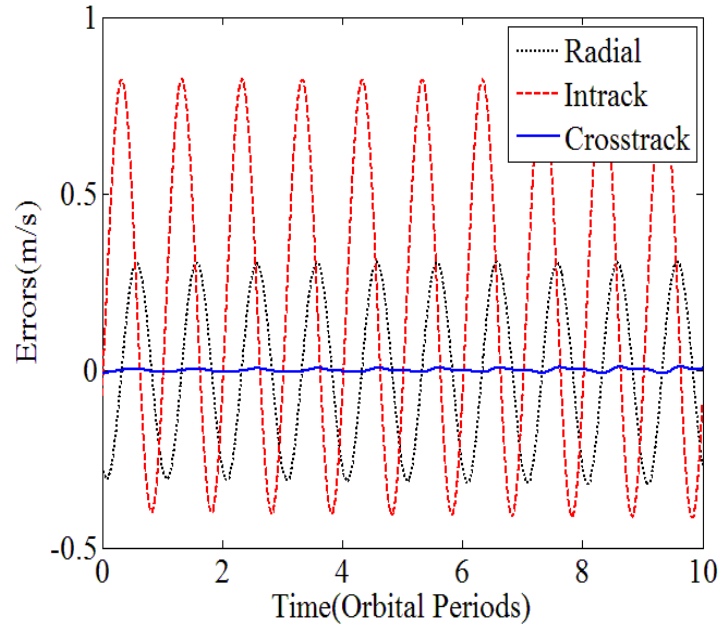


Figure 3-30: Time history of velocity errors for HCW equations without using estimated initial conditions.

3.7.2 Orbit Eccentricity

HCW equations are derived for circular orbit of chief with $e = 0$. Therefore, its solutions are not valid for moderate or highly eccentric orbits. Figures: 3-31 and 3-32 shows how different eccentricities of chief orbit effect errors in LVLH frame between reference and HCW modelled relative motion. The chief satellite is still with same initial conditions as for SAR-Lupe 1. However, one now varies the range of eccentricities for this satellite and estimate the orbit of deputy satellite for one orbital period to find out the growth of errors over time. A criterion for maximum errors of 5% between the true relative and linearized (HCW) relative motion in LVLH coordinates of free ellipse orbital size is set. The idea is to observe different eccentricities of chief orbit for this measure to hold good. The range of eccentricities comes out to be $(0 \leq e \leq 2.5 \times 10^{-3})$. Therefore, the choice of “ e ” can be made depending on the maximum allowable error.

3.7.3 Semi Major Axis and Inclination

The orbit semi major axis and inclination are two parameters which appear in the expressions (Equation: 3.36) for secular and periodic terms in J_2 perturbed orbits. Since, HCW equations assume spherically symmetric geopotential for absolute motion of chief satellite therefore; changing these parameters will impact differences in true and linearized (HCW) relative motion. Similarly to the analysis in Section: 3.7.2, one again sets the criterion for maximum errors between the true relative and linearized (HCW) relative motion as 5% in LVLH coordinates of free ellipse orbital size. The

errors are observed to be less than 5% of free ellipse orbital size in all the three directions over one orbital period (shown in Figures: 3.33 to 3.35).

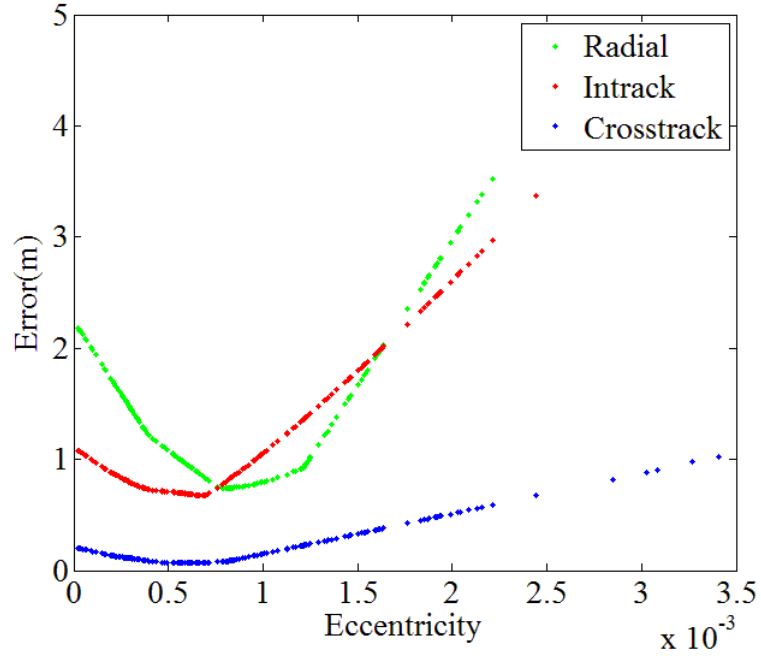


Figure 3-31: Maximum position errors for HCW equations with optimal initial conditions $\hat{\mathbf{x}}_0$ over 1 orbital period.

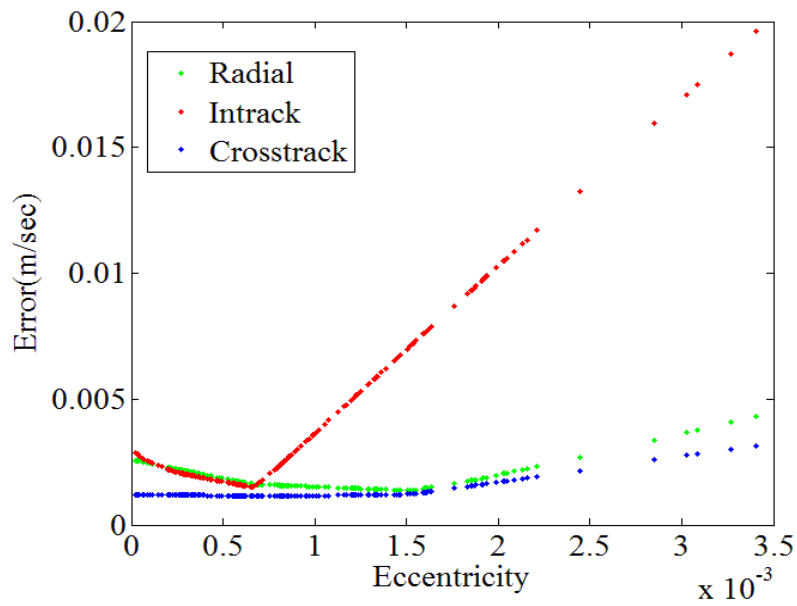


Figure 3-32: Maximum velocity errors for HCW equations with optimal initial conditions $\hat{\mathbf{x}}_0$ over 1 orbital period.

However, there is an increased positional error at lower inclination and lower orbital semi major axis due to a more pronounced effect of oblateness of Earth near equatorial inclinations and reduced distance from the main gravitational force i.e., Earth, respectively. As IC are chosen by minimizing the variance, therefore the error growth is significantly less compared to error statistics provided in Table: 3-8 for initial conditions selected without estimation.

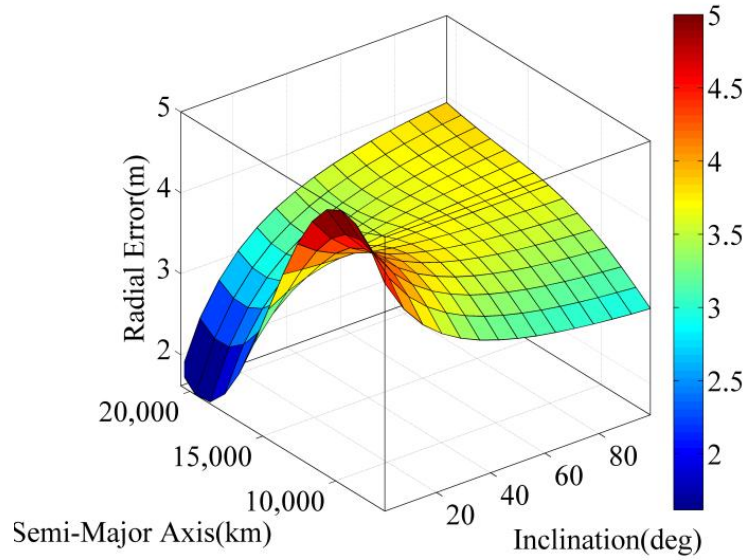


Figure 3-33: Maximum position errors (radial direction) for HCW model over one orbital period using optimal IC.

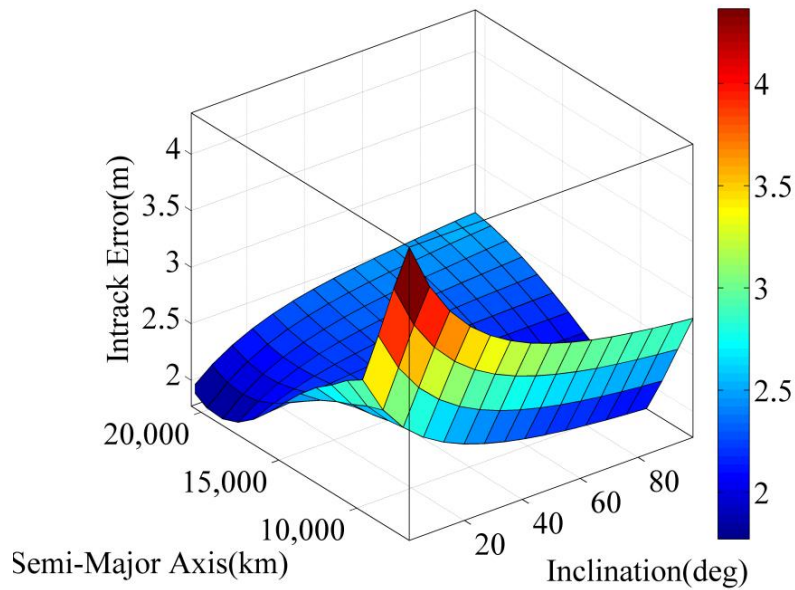


Figure 3-34: Maximum position errors (in-track direction) for HCW model over one orbital period using optimal IC.

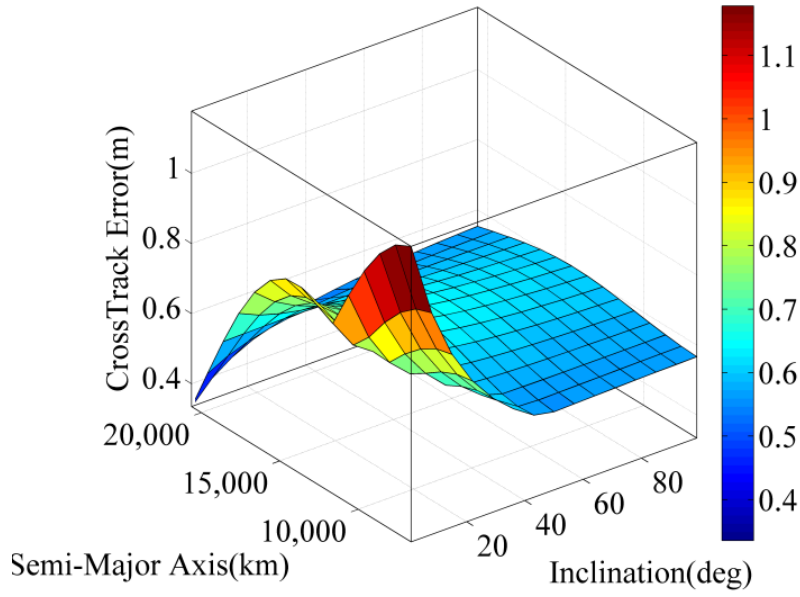


Figure 3-35: Maximum position errors (cross-track direction) for HCW model over orbital period using optimal IC.

3.7.4 J_2 Modified HCW Equations by Schweighart and Sedwick

Analysis of HCW equations in Sections: 3.7.1 to 3.7.3 reveals that its solution would break down if the assumptions of circular chief orbit, spherically symmetric geopotential and linearized differential gravity accelerations are violated (see Figure: 3-29 and 3-30). Therefore, a need was felt to derive equations that describe the relative motion of satellites under the influence of eccentric chief orbit, non linear differential gravity and oblate Earth. The zonal spherical harmonic J_2 , due to oblate Earth, being the most significant perturbation will be considered here for analysis of the relative motion of satellites. In this section a modification of the HCW equations for J_2 perturbed relative motion of satellites given by SS [20] will be analyzed. The procedure for analysis of HCW equations will now be repeated for SS model. The J_2 modified HCW equations for relative motion between two satellites under the effect of J_2 is given by three second order linear differential equations expressed as under [20]:

$$\begin{aligned}\ddot{x} - 2nc\dot{y} - (5c^2 - 2)n^2x &= 0 \\ \ddot{y} + 2nc\dot{x} &= 0 \\ \ddot{z} + q^2z &= 2lq \cos(qt + \phi)\end{aligned}\tag{3.59}$$

where,

$$c = \sqrt{1+s}, \quad n = \sqrt{\frac{\mu_E}{r_c^3}}, \quad s = \frac{3J_2 R_E^2}{8r_c^2} (1 + 3\cos 2i_c),$$

$$q = nc - (\cos \gamma_0 \sin \gamma_0 \cot \Delta\Omega_0 - \sin^2 \gamma_0 \cos i_d)$$

$$\gamma_0 = \cot^{-1} \left[\frac{\cot i_c \sin i_d - \cos i_d \cos \Delta\Omega_0}{\sin \Delta\Omega_0} \right], \quad \Delta\Omega_0 = \frac{z_0}{r_c \sin i_c}$$

r_c = chief orbital radius

i_c = inclination of chief orbit

i_d = inclination of deputy orbit

In this model the angular velocity vector ($\vec{\omega}_c$) of rotating frame is slightly modified. The value of mean motion “ n ” ($n = |\vec{\omega}|$) is slightly varied by a factor “ c ” for in-plane (x - y) motion. Equations are still coupled in in-plane, and decoupled in out-of-plane directions. Moreover, angular frequency of cross track motion is changed to “ q ”. The analytic solutions to modified HCW equations (Equation: 3.59) as found out by SS are presented in Appendix-E [20]. The analytical solutions of SS relative motion model are firstly transformed from LVLH into ECI coordinate frame, following the methodology of Equations: 3.54-3.55 and will not be repeated here for the sake of clarity. However, the partial differential matrix for SS model in the estimation problem is presented in Appendix-F. On similar lines to HCW equations, orbital data for the reference deputy satellite is obtained by numerical integration of the nonlinear equation of motion (Equation: 3.32). The orbital data for linear approximation of the deputy satellite is obtained using the transformed SS model (transformed from LVLH to ECI). The estimation process for J_2 Modified HCW Equations is now carried out. The initial conditions provided to the deputy satellite in LVLH coordinate frame obtained from the SS model with $A = 50$ m and $\phi = 65$ deg [20] and estimated IC are given in Table: 3-9.

IC in LVLH Coordinates frame	IC of Numerical Trajectory \mathbf{x}_0	IC of Analytical Trajectory (Output of estimator - $\hat{\mathbf{x}}_0$)
x_0	$A \cos \phi = 21.13$ m	12.20 m
y_0	$2A \sin \phi = 90.63$ m	90.37 m
z_0	$2A \sin \phi = 90.63$ m	90.75 m
\dot{x}_0	0.0503 m/s	0.3127 m/s
\dot{y}_0	-0.0469 m/s	-0.0159 m/s
\dot{z}_0	0.0469 m/s	0.1111 m/s

Table 3-9: Comparison of IC for numerical and analytical trajectory for HCW equations compared with J_2 perturbed full nonlinear relative motion equations.

Non-zero velocity terms for (x) and (y) (see Table: 3-9) for the numerical trajectory are used to remove the drift and offset terms of the SS solution (see Appendix-E for details). Errors in the relative orbit of the deputy satellite propagated by numerical integration of equations of motion and analytical SS model with the newly estimated state for 10 orbital periods is shown in Figures: 3-36 and 3-37. Table: 3-10 summarizes the position and velocity error statistics. A considerable improvement is observed in all the three LVLH coordinates. Although, in-track errors are comparatively larger than others, however, they provide useful insight into the dynamics and their growth is smaller when compared with the HCW model (Figure: 3.27). Firstly because of the modification carried out in the mean motion of the chief satellite by a factor “ c ” in in-plane (x - y plane). Secondly, the drift rate has also been reduced due to correct initial conditions applied for elimination of the secular growth in in-track motion (given in Appendix-E). Moreover, the in-plane growths are periodic in nature. In radial direction the error is about max 10 m. In cross-track the error is max 5 m. For comparison purposes the initial conditions \mathbf{x}_0 (Table: 3-9) are now used to observe the deputy satellite without estimation. The error plots for position and velocity are shown in Figures: 3-38 and 3-39. As expected, the errors are substantial, especially in in-track direction where it grows up to 2.744 km in 10 orbital periods; whereas, the cross track error and radial errors are 260 m and 7.4 m, respectively.

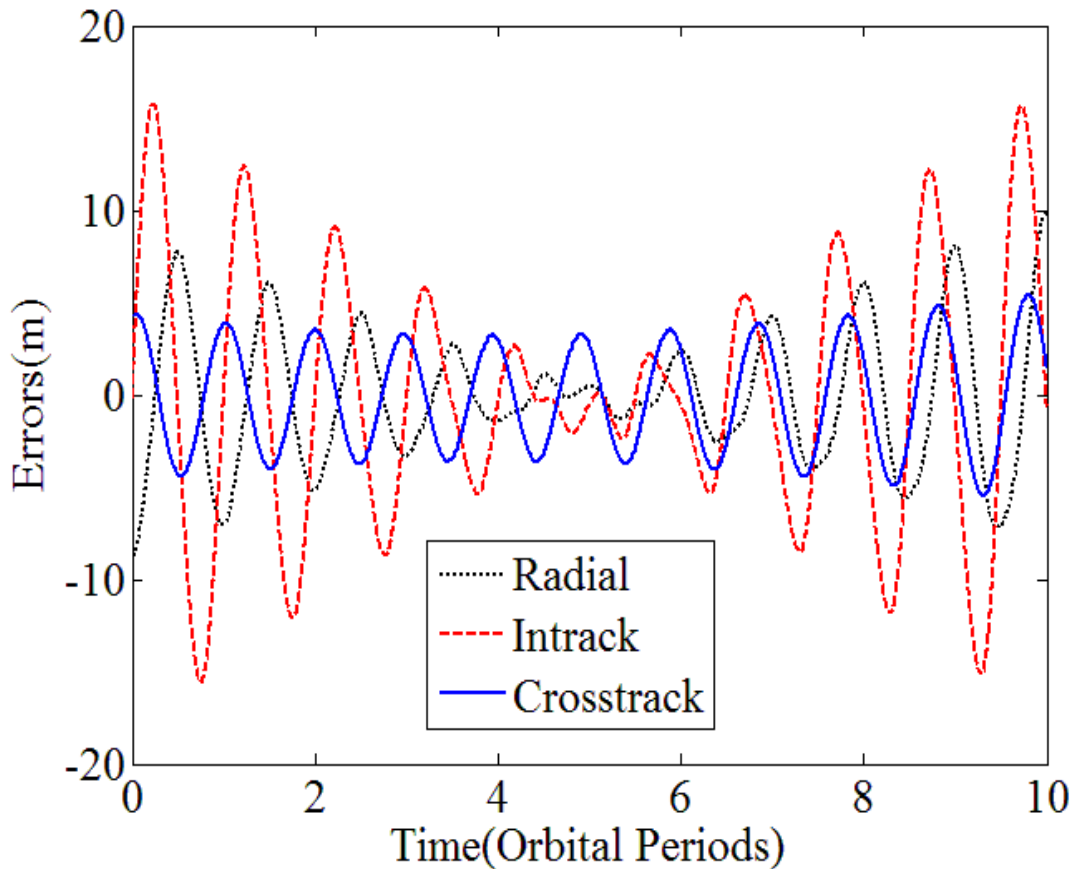


Figure 3-36: Time history of position errors for SS model after using optimal initial conditions.

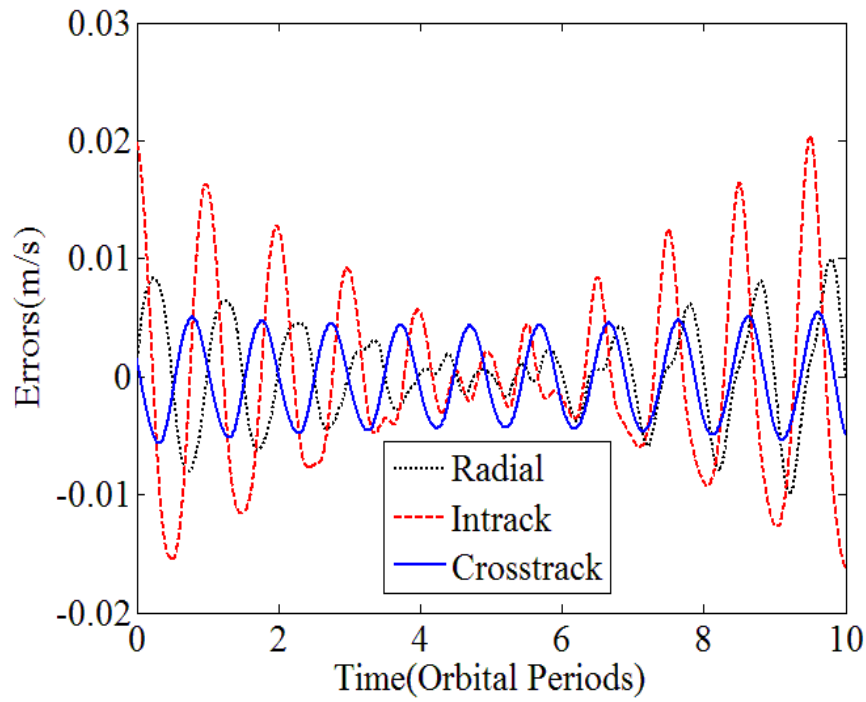


Figure 3-37: Time history of velocity errors for SS model after using optimal initial conditions.

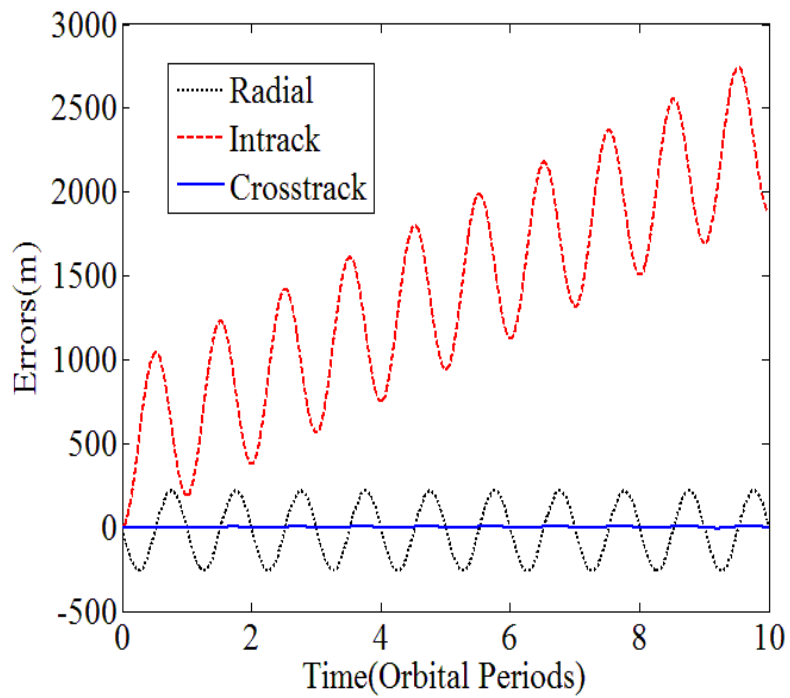


Figure 3-38: Time history of position errors for SS model without modifying initial conditions.

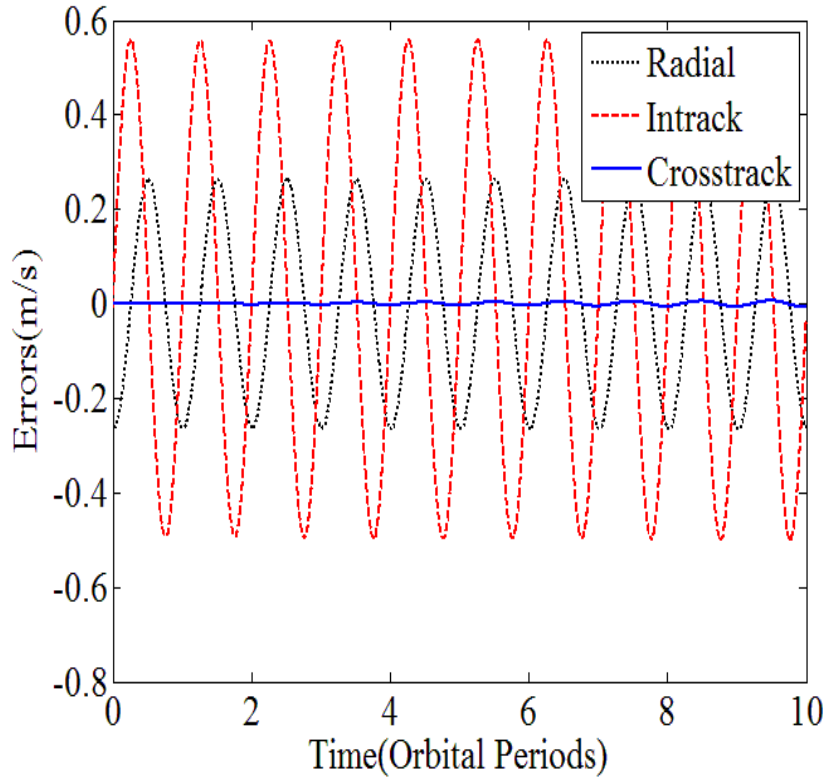


Figure 3-39: Time history of velocity errors for SS model without modifying initial conditions.

IC used for propagation of SS equations	Position (m)			Velocity (m/s)		
	R	I	C	R	I	C
\mathbf{x}_0	260	2744	7.4	0.26	0.56	0.007
$\hat{\mathbf{x}}_0$	10	16	5	0.010	0.020	0.005
R – Radial, I – In-track, C – Cross-track \mathbf{x}_0 – IC of numerical trajectory $\hat{\mathbf{x}}_0$ – Estimated IC (output of estimator)						

Table 3-10: Summary of Maximum Absolute Position and Velocity Errors in LVLH Coordinates over 10 Orbital Periods for SS Model

3.7.5 Conclusion

In this section a comparison of two relative motion models for satellite flight formations was carried out. Firstly, the HCW model has been investigated with a view to analyze its process noise and assessment of its fidelity compared with the reference nonlinear relative motion model. The choice of initial condition is very critical to minimize effects of un-modelled accelerations and nonlinearity of true equations of motion. Using appropriate IC, the errors in the HCW model are reduced to 95.3% in

radial, 99.68% in in-track and 42% in cross track directions, respectively. The analysis on different choices of the semi-major axis, inclination and eccentricity of chief orbit has also been undertaken. The choice of appropriate IC is essential to reduce the unbounded error growth for all such choices. Secondly, the SS relative motion model is analyzed for its process noise and evaluation of fidelity. SS model provides improvement of HCW equations for J_2 perturbed relative motion. However, due to averaging of second zonal harmonic J_2 over entire orbital period [20], could not truly capture the true secular and periodic variations in the orbit. Nevertheless, by using appropriate IC the reduction in LVLH errors amounts to 96% in radial, 99.4% in in-track and 32.4% in cross track.

3.8 Free Propagation Error Growth

The orbital models discussed in this analysis will now be assessed for growth of error in LVLH position coordinates when propagated forward in time after being initialized with optimal initial conditions. The choice of such of initial conditions is based on estimation for only one orbital period. The error of 10 m in any direction i.e., radial, in-track and cross-track is selected as maximum allowable during the forward propagation. Essentially one would observe the time for which the errors for particular orbital model remain bounded inside a cube with 10 m on each side. Table: 3-11 provide the error statistics for positions. The plot for HCW, SS and Epicycle model are shown in Figures: 3-40 to 3-42.

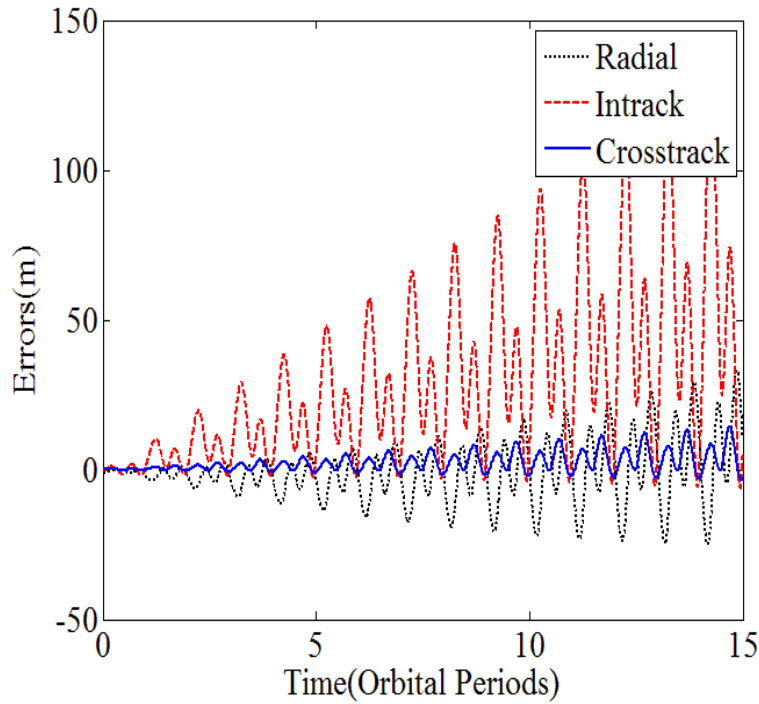


Figure 3-40: Time history of growth of position errors for HCW model using optimal initial conditions (based on estimation for one orbital period).

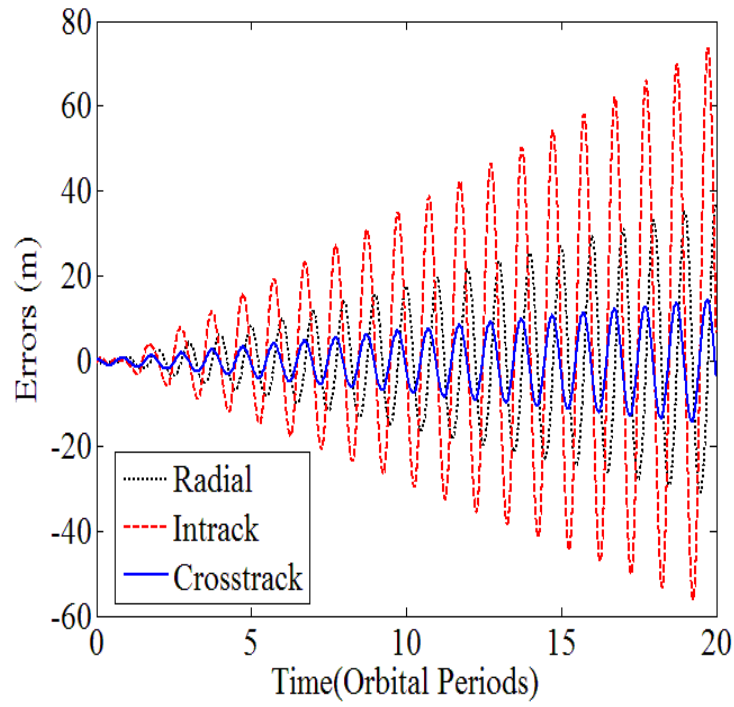


Figure 3-41: Time history of growth of position errors for SS model using optimal initial conditions (based on estimation for one orbital period).

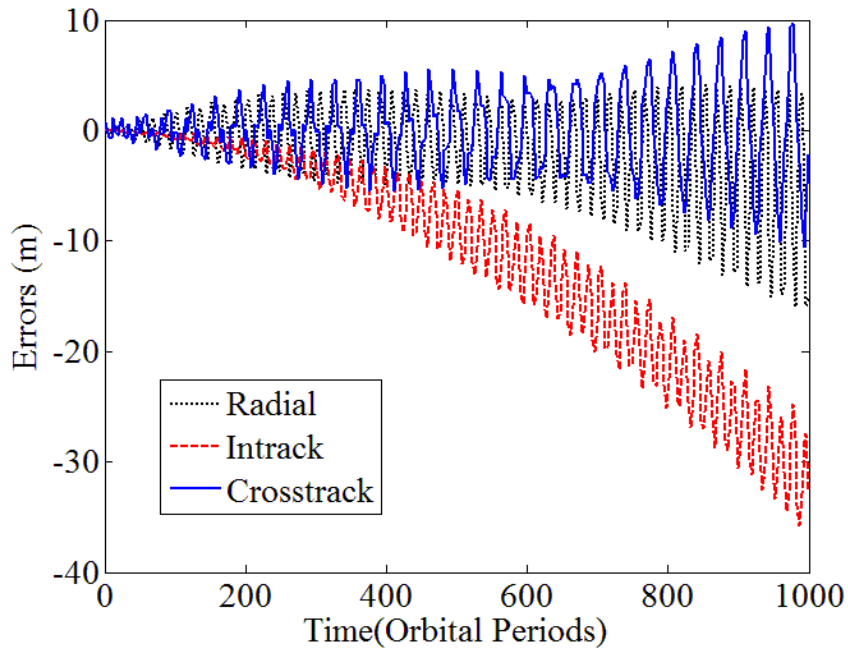


Figure 3-42: Time history of growth of position errors for Epicycle model using optimal initial conditions (based on estimation for one orbital period).

The error for HCW and SS model reaches 10 m, well before one day for a deputy satellite selected in

this example with chief orbit as of SAR Lupe-1. However, for epicycle model the errors remain bounded for 50.3 days (766 orbital periods) in radial, 29.07 days (442 orbital periods) in in-track, and 62.96 (957 orbital periods) days in cross-track directions, respectively.

Model for analytical trajectory	Time (Orbital Periods) when error exceeds 10 m (in any direction) in LVLH coordinates frame		
	R	I	C
HCW	4.10	1.24	10.69
SS	6	3.72	15
Epicycle	766	442	957
R – Radial, I – In-track, C – Cross-track			

Table 3-11: Summary of absolute position errors limit criteria of 10 m in any three directions of LVLH coordinates frame.

3.9 Summary

In this chapter an analysis of fidelities of different orbital models has been carried out. The error statistics are tabulated to assess their long term growth. The analytical models greatly enhance understanding of complex orbital motion. However, due to simplifications and neglecting the actual dynamics may lead to considerable errors especially for formation flying missions. The initial conditions found out through estimation do not produce the desired absolute or relative motion. They are meant to generate analytical model satellite trajectories which are very close to the reference nonlinear trajectory produced by numerical integration. By using adapted GLDC estimator formulations, considerable reduction in positional errors could be achieved in all the three directions. Furthermore, by using estimated IC one is able to compare the validity and usefulness of analytical models over a period of time (see Table: 3-11). This has implications on use of a particular analytic model for close orbiting satellites. For example, consider a satellite formation with inter satellite distance of < 50 m. In this scenario new estimate of IC would be required more frequently (less than a day) for HCW or SS models compared to epicycle model which provides months of accuracy without updating orbital parameters. These formulations can be also used to generate forward propagation for evolution of orbits in sequential state estimators and orbit controls. In orbit control scenario one may modify the initial conditions to generate an orbit which is very close to perturbed orbit and then apply control corrections to achieve the desired trajectory.

4 Epicycle Orbit Parameter Filter

4.1 Introduction

In Section: 3.4 orbital dynamics of a satellite around a *non-spherical* geopotential were described. When the perturbing forces are conservative, as with the gravitational perturbation due to non-spherical nature of our planet, then the accelerations (Equation: 3.32) are expressed as gradients of the disturbing function (Equation: 3.30). Axisymmetric geopotential i.e., symmetric about the North Pole of the Earth will be considered here. In general, numerical integration of equations of motion of a satellite in a *non-spherical* gravitational field would yield its high precision ephemerides. Preceding analysis of Chapter: 3, focused on mathematical models for LEO satellites influenced by accelerations due to spherical gravitational potential and perturbations due to J_2 . There a methodology of fitting an approximate model to nonlinear data by adapting GLDC scheme (Section: 3.2) was discussed. The high fidelity nature of analytic epicycle model [2] was clearly evident especially due to the lower error growth i.e., few meters over long durations (see Table: 3-11) compared to other analytical models like Kepler's equation [13] (see Table: 3-4), HCW [18],[19] and SS model [20] (see Table: 3-10). The epicycle model is capable of describing all the gravitational perturbative effects arising due to the oblate shape of the Earth. The higher order zonal effects can be incorporated in terms of coefficients for secular, long and short periodic variations in the orbit (see Equation: 3.35). These higher order zonal effects can be further expressed in terms of even and odd harmonic for variations in the orbit. There are no long periodic and secular variations in the orbit due to even and odd zonal harmonics, respectively [2]. Therefore, we denote the higher *even harmonics* with subscript $2m$, and the terms due to *odd harmonics* with subscript $2m+1$. The coefficient for radial offset (ρ), coefficient for secular variation in RAAN (ϑ), and coefficient for secular variation in argument of latitude (κ) in Equation: 3.35 can be extended as:

$$\rho = \sum_{m=1}^M \rho_{2m} + \rho_{2m+1}, \quad \vartheta = \sum_{m=1}^M \vartheta_{2m} + \vartheta_{2m+1}, \quad \kappa = \sum_{m=1}^M \kappa_{2m} + \kappa_{2m+1} \quad (4.1)$$

The coefficient inside summation are calculated as [2],[14]:

$$\rho_{2m} = (2m - 1)A_{2m}L_{2m}^0 \quad (4.2)$$

$$\vartheta_{2m} = -2A_{2m} \frac{\cot I_0}{\sin I_0} \sum_{l=1}^m 2l L_{2m}^{2l}$$

$$\kappa_{2m} = -\vartheta_{2m} \cos I_0 - (4m - 1)A_{2m}L_{2m}^0$$

where,

$$A_{2m} = J_{2m} \left(\frac{R_E}{a} \right)^{2m} \quad (4.3)$$

$$L_i^j = \frac{(i-j)!}{(i+j)!} P_i^j(0) P_i^j(\cos I_0)$$

Note that the order $M = \infty$. However for practical reasons, yet sufficiently accurate requirements the series can be truncated up to a certain order for example $M = 20$ for 41×41 WGS-84 model [12] and Legendre function $P_i^j(x)$ is defined as:

$$P_i^j(x) = (1-x^2)^{j/2} \frac{d^j}{dx^j} P_i(x) \quad (4.4)$$

As there are no secular variations in the orbit due to odd zonal harmonics therefore [2]:

$$\rho_{2m+1} = \vartheta_{2m+1} = \kappa_{2m+1} = 0 \quad (4.5)$$

For short periodic coefficients Δ_x (see Equation: 3.35) one has the expression

$$\begin{aligned} \Delta_r &= \sum_{m=1}^M \Delta_{r_{2m}} + \Delta_{r_{2m+1}}, & \Delta_I &= \sum_{m=1}^M \Delta_{I_{2m}} + \Delta_{I_{2m+1}} \\ \Delta_\Omega &= \sum_{m=1}^M \Delta_{\Omega_{2m}} + \Delta_{\Omega_{2m+1}}, & \Delta_\lambda &= \sum_{m=1}^M \Delta_{\lambda_{2m}} + \Delta_{\lambda_{2m+1}} \end{aligned} \quad (4.6)$$

The even short periodic perturbations inside summation of Equation: 4.6 can be expressed as:

$$\begin{aligned}\Delta_{r_{2m}} &= \sum_{l=1}^m \Delta r_{2m}^l \cos 2l\alpha, & \Delta_{I_{2m}} &= \sum_{l=1}^m I_{2m}^l (1 - \cos 2l\alpha) \\ \Delta_{\Omega_{2m}} &= \sum_{l=1}^m \Delta \Omega_{2m}^l \sin 2l\alpha, & \Delta_{\lambda_{2m}} &= \sum_{l=1}^m \Delta \lambda_{2m}^l \sin 2l\alpha\end{aligned}\tag{4.7}$$

where,

$$\begin{aligned}\Delta r_{2m}^l &= 2(2m-1)A_{2m} \frac{(-1)^l}{1-(2l)^2} L_{2m}^{2l} \\ I_{2m}^l &= 2A_{2m}(-1)^l \cot I_0 L_{2m}^{2l} \\ \Delta \Omega_{2m}^l &= -2A_{2m}(-1)^l \frac{\cot I_0}{\sin I_0} \Lambda_{2m}^{2l} \\ \Delta \lambda_{2m}^l &= -\Delta \Omega_{2m}^l \cos I_0 - 2A_{2m} \frac{(-1)^l}{2l} \left[1 + \frac{2(2m-1)}{1-(2l)^2} \right] L_{2m}^{2l} \\ \Lambda_{2m}^{2l} &= L_{2m}^{2l} + 2 \sum_{k=l+1}^m \frac{2k}{2l} L_{2m}^{2k}\end{aligned}\tag{4.8}$$

Finally the short periodic variations due to odd zonal harmonics are [14]:

$$\begin{aligned}\Delta_{r_{2m+1}} &= \sum_{l=1}^m \Delta r_{2m+1}^l \sin(2l+1)\alpha, \\ \Delta_{I_{2m+1}} &= \sum_{l=1}^m I_{2m+1}^l \sin(2l+1)\alpha \\ \Delta_{\Omega_{2m+1}} &= \sum_{l=1}^m \Delta \Omega_{2m+1}^l [1 - \cos(2l+1)\alpha] \\ \Delta_{\lambda_{2m+1}} &= \sum_{l=1}^m \Delta \lambda_{2m+1}^l [1 - \cos(2l+1)\alpha]\end{aligned}\tag{4.9}$$

where,

$$\begin{aligned}
\Delta r_{2m+1}^l &= 4mA_{2m+1} \frac{(-1)^l}{1 - (2l+1)^2} L_{2m+1}^{2l+1} \\
\Delta I_{2m+1}^l &= -2A_{2m+1}(-1)^l \cot I_0 L_{2m+1}^{2l+1} \\
\Delta \Omega_{2m+1}^l &= -2A_{2m+1}(-1)^l \frac{\cot I_0}{\sin I_0} \Lambda_{2m+1}^{2l+1} \\
\Delta \lambda_{2m+1}^l &= -\Delta \Omega_{2m+1}^l \cos I_0 - 2A_{2m+1} \frac{(-1)^l}{2l+1} \left[1 + \frac{4m}{1 - (2l+1)^2} \right] L_{2m+1}^{2l+1} \\
\Lambda_{2m+1}^{2l+1} &= L_{2m+1}^{2l+1} + 2 \sum_{k=l+1}^m \frac{2k+1}{2l+1} L_{2m+1}^{2k+1}
\end{aligned} \tag{4.10}$$

Now the long periodic variations in radial coordinates (first of Equation: 3.35) is expressed as [2]:

$$\begin{aligned}
\chi &= \sum_{m=1}^M \chi_{2m} + \chi_{2m+1} \\
\chi_{2m+1} &= - \left(2m \frac{A_{2m+1}}{\kappa_2} \right) L_{2m+1}^1
\end{aligned} \tag{4.11}$$

Since there are no long periodic variations in the orbit due to even zonal harmonics therefore:

$$\chi_{2m} = 0 \tag{4.12}$$

All the perturbative effects due to zonal harmonics (Equation: 4.1 to 4.12) can be conveniently incorporated into the epicycle evolution equation (Equation: 3.35) to obtain a high precision analytical trajectory. However, being an analytical orbital description its long term propagation needs a proper choice of the orbital parameters.

In the preceding analysis (Chapter: 3), it was observed that by carefully choosing the orbital parameters for an analytical approximation appropriate to a given choice of parameters for the numerically propagated orbit obtained from nonlinear equations of motion would keep the two trajectories sufficiently close to each other for long times and can minimize error growths (see Figure: 3-2). However, one of the main difficulties / complexities associated with methodology of Section: 3.2 is the calculation of the partial derivative matrix \mathbf{F} (see Equation: 3.9 and Appendix-D) for use in estimation of these parameters. The complexity would further enhance if one extends the perturbative terms to higher order i.e., greater than two. Therefore, in order to generate a higher order analytical trajectory an alternate methodology is adopted for estimation of epicycle orbital parameters. The method exploits the linear secular nature of epicycle coordinates of argument of latitude “ λ ” and right ascension of the ascending node “ Ω ” (RAAN) (see Equation 3.35 and Figure: 3-19). The new

parameter estimation technique is developed using the linear least squares method [39],[46] and is called Epicycle Parameter Filter (EPF). The fundamental idea is minimization of the process noise in the epicycle model in order to enhance its validity over long periods of time when compared with full nonlinear equations of motion. Here, few definitions for an epicyclic orbit would be recalled. The position of a satellite in an epicyclic orbit is defined by six osculating coordinates (Equation: 3.35). The geometrical shape of an epicyclic orbit is described by six constant parameters; semi-major axis (a), inclination (I_0), right ascension of ascending node (Ω_0), non singular parameters for undefined epicycle phase at perigee passage (needed for equatorial orbits) (ξ_P, η_P) and an equator crossing time (t_E). Extremely precise selection of these parameters is needed to obtain epicyclic orbital coordinates appropriate to a given numerically propagated orbital coordinates. Similarly to Chapter: 3, one would use the term “reference or numerical trajectory” for nonlinear orbital data obtained from numerical integration of equations of motion and analytical trajectory for linearized orbital data obtained from analytic epicyclic equations (Equation: 3.35).

4.2 Secular Variations in Epicycle Orbital Coordinates

The expressions for the argument of latitude, λ and RAAN, Ω in Equation: 3.35 for epicycle coordinates contain secularly growing linear quantities (see Figures: 4.1 and 4.2) depending on coefficients of κ and ϑ . The equations for these coordinates are expressed as [2]:

$$\begin{aligned}\lambda &= \beta + 2\xi_P \sin \alpha + 2\eta_P(1 - \cos \alpha) - 2\chi(1 - \cos \beta) + \Delta\lambda_2 \sin 2\beta + \Delta\lambda \\ \Omega &= \Omega_0 + \vartheta\alpha + \Delta\Omega_2 \sin 2\beta + \Delta\Omega\end{aligned}\tag{4.13}$$

where,

$$\beta = (1 + \kappa)\alpha,$$

$\Delta\lambda_2$ = short periodic coefficients due to J_2

$\Delta\lambda$ = higher order short periodic variations, derived from Equation: 4.6

χ = higher order long periodic coefficient derived from Equation: 4.11

Equations: 4.2 and 4.3 reveals the dependence of the coefficients of secular change (κ and ϑ) on semi-major axis a , and inclination I_0 . Note the secular growth in epicycle coordinates is significantly more dominant than periodic variations. Therefore, one has to accurately fix coefficients of κ and ϑ in epicyclic evolution equations (Equation: 3.35) in order to obtain high precision long term (i.e., weeks) secular variations in the orbit when compared with the coordinates of the numerical trajectory.

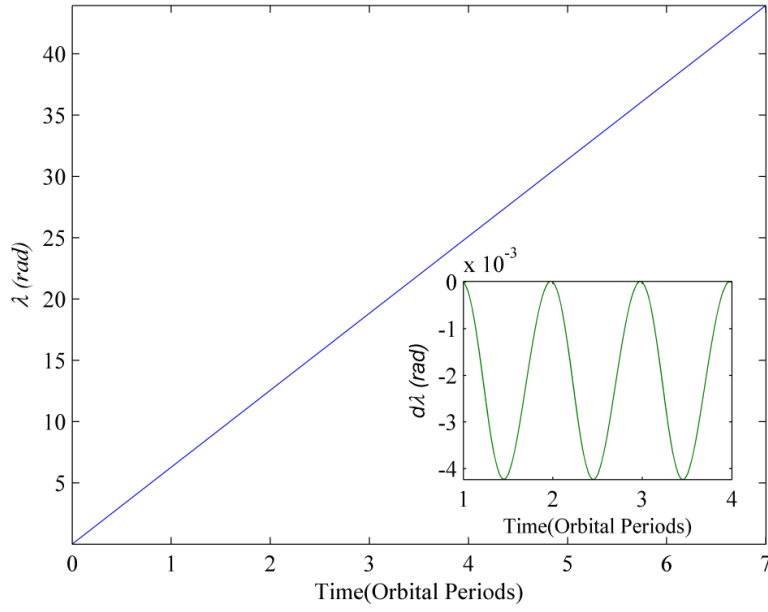


Figure 4-1: The above plot depicts the dominant linear secular growth and small periodic variations in λ . The inner plot (shown in green) is an augmented view to observe the oscillating terms, in Equation: 4.13 (see Equation: 4.20 for description of $d\lambda$), which are otherwise not viewable in main (shown as straight blue line).

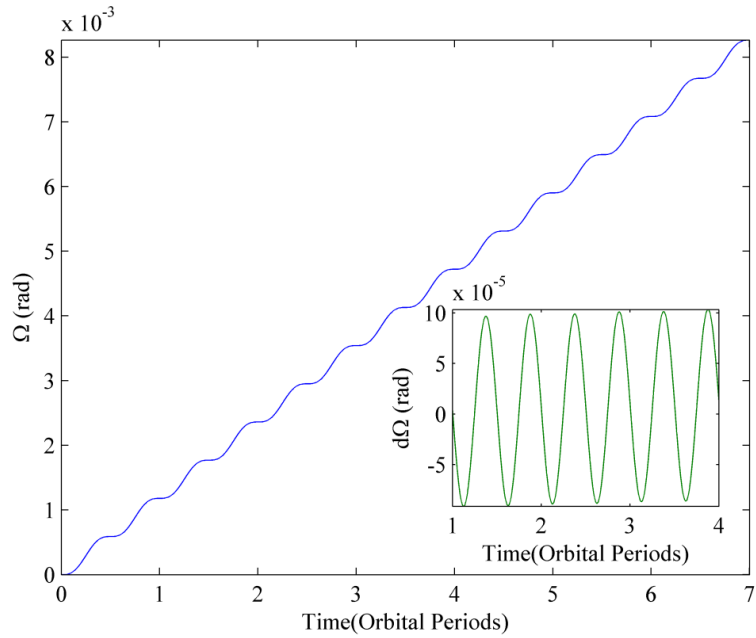


Figure 4-2: The above plot depicts the dominant linear secular growth and small periodic variations in Ω . The inner plot (shown in green) is an augmented view to observe the oscillating terms, in second of Equation: 4.13 (see Equation: 4.20 for description of $d\Omega$).

4.3 Development of an Epicycle Parameter Filter

The ensuing description develops an orbital parameter estimation algorithm called as EPF by using statistical data regression technique known as linear least squares [5]. The epicycle parameters to be estimated are:

$$\mathbf{x}_0 = (a, \xi_P, \eta_P, I_0, \Omega_0, t_E)^T \quad (4.14)$$

4.3.1 Reference Nonlinear Satellite Trajectory

The reference nonlinear orbital dynamical equations in ECI coordinate frame for the non-spherical geopotential are expressed as [13]:

$$\dot{\mathbf{r}} = \mathbf{v} \quad (4.15)$$

$$\dot{\mathbf{v}} = -\frac{\mu_E}{r^3} \mathbf{r} + \mathbf{a}_G$$

where, $\mathbf{r} = (X, Y, Z)^T$, is the position vector, $\mathbf{v} = (\dot{X}, \dot{Y}, \dot{Z})^T$, is the velocity vector, both in ECI coordinates, and \mathbf{a}_G expresses as higher zonal gravitational perturbation terms obtained by taking the gradient of potential function “ \mathcal{U} ” given in Equation: 3.30.

The prediction of precise satellite ephemerides are obtained by numerically integrating these equations given some epoch satellite state $(\mathbf{r}_0, \mathbf{v}_0)^T$.

The nonlinear orbital data required by the EPF is in terms of epicycle orbital coordinates; whereas, its availability is in terms of ECI coordinate frame (Equation: 4.15). Therefore, the first step is to numerically integrate Equation: 4.15 including zonal harmonic perturbation terms up to a certain order for a specific duration i.e., a week. This numerical trajectory describes the satellite’s position and velocity in a three dimensional ECI coordinate system at specific instants of time, from epoch time t_0 to some later time t_k expressed as:

$$\mathbf{Z}(\mathbf{x}_0, t) = [(\mathbf{r}_0, \mathbf{v}_0), (\mathbf{r}_1, \mathbf{v}_1), \dots (\mathbf{r}_k, \mathbf{v}_k)]^T \quad (4.16)$$

where, k is the time subscript for state vectors of a satellite. A transformation is applied to convert

each position and velocity vector into epicycle coordinates by using [38],[12]:

$$\begin{aligned}
 r &= \sqrt{X^2 + Y^2 + Z^2}, \quad \mathbf{v} = \sqrt{\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2} \\
 \lambda &= \cos^{-1} \left(\frac{\dot{Z} - \frac{v_r Z}{r}}{v_\theta \sin I} \right), \quad \text{if } \begin{cases} (Z/(r \sin I) < 0), & \lambda = 2\pi - \lambda \\ (\lambda > \pi) & \lambda = \lambda - 2\pi \end{cases} \\
 I &= \cos^{-1} \left(\frac{X\dot{Y} - Y\dot{X}}{rv_\theta} \right) \\
 \Omega &= \cos^{-1} \left(\frac{X\dot{Z} - Z\dot{X}}{rv_\theta \sin I} \right), \quad \text{if } \left\{ \left((Y\dot{Z} - Z\dot{Y}) < 0 \right), \quad \Omega = 2\pi - \Omega \right. \\
 v_r &= \frac{\mathbf{r} \cdot \mathbf{v}}{r}, \quad v_\theta = \sqrt{v^2 - v_r^2}
 \end{aligned} \tag{4.17}$$

The transformed data from Equation: 4.17 would be termed as “nonlinear or numerical trajectory” consisting of satellite nonlinear epicycle coordinates:

$$\begin{aligned}
 \mathbf{Z}_{Epicycle}(\mathbf{x}_0, t) &= [(r(t_0), \lambda(t_0), I(t_0), \Omega(t_0), v_r(t_0), v_\theta(t_0)) \dots \\
 &\dots (r(t_k), \lambda(t_k), I(t_k), \Omega(t_k), v_r(t_k), v_\theta(t_k))]^T
 \end{aligned} \tag{4.18}$$

where, $\mathbf{Z}_{Epicycle}(\mathbf{x}_0, t)$ is the vector of nonlinear epicycle trajectory coordinates.

4.3.2 Least Squares Formulation

Coordinates of argument of latitude (λ) and RAAN (Ω) are angular descriptions repeating themselves after an orbital period. Therefore, data for these coordinates from Equation: 4.18 are unravelled to obtain time evolution of continuously increasing angular quantities. One would unravel λ and Ω so that these grow linearly instead of the usual $(-\pi \leq \lambda < \pi)$ and $(0 \leq \Omega < 2\pi)$, respectively. The equation for these two coordinates may be separated into linear and oscillating terms therefore one may rewrite the terms from Equation: 4.13:

$$\begin{aligned}
 \tilde{\lambda} &= \lambda - d\lambda \triangleq \beta \\
 \tilde{\Omega} &= \Omega - d\Omega \triangleq \Omega_0 + \vartheta\alpha
 \end{aligned} \tag{4.19}$$

where,

$$\begin{aligned}
d\lambda &= 2\xi_P \sin \alpha + 2\eta_P(1 - \cos \alpha) - 2\chi(1 - \cos \beta) + \Delta\lambda_2 \sin 2\beta + \Delta\lambda \\
d\Omega &= \Delta\Omega_2 \sin 2\beta + \Delta\Omega
\end{aligned} \tag{4.20}$$

Equation: 4.19 can be rewritten as following:

$$\begin{aligned}
\tilde{\lambda} &= (1 + \kappa)n(t - t_E) = p_1(t - t_E) \\
\tilde{\Omega} &= \Omega_0 + \vartheta n(t - t_E) = \Omega_0 + p_2(t - t_E)
\end{aligned} \tag{4.21}$$

where, p_1 and p_2 are slopes of λ and Ω for the numerical trajectory, respectively; whereas, $(1 + \kappa)n$ and ϑn are slopes of these coordinates for the analytical trajectory (see Figure: 4.1 and 4.2) and $n = \sqrt{\mu_E/a^3}$ (mean motion).

As $\tilde{\lambda}$ and $\tilde{\Omega}$ are linearly increasing coordinates therefore, the linear least squares method can be used to estimate p_1, p_2, Ω_0 and t_E . The reason for this assumption is quite valid as linear secular growth in angular quantities of λ and Ω are more dominant than oscillating terms, $d\lambda$ and $d\Omega$ (see Figure: 4.1 and 4.2). Therefore, these do not make much impact on the linear least square fit. The cost function for linear least squares problem can now be conveniently written as:

$$\mathcal{J} = \sum_{i=0}^k [\tilde{\lambda}(t_i) - p_1(t_i - t_E)]^2 + \sum_{i=0}^k [\tilde{\Omega}(t_i) - \Omega_0 - p_2(t_i - t_E)]^2 \tag{4.22}$$

The cost function is now differentiated with respect to four variables i.e., p_1, p_2, Ω_0 and t_E and equated to zero.

$$\frac{\partial \mathcal{J}}{\partial (p_1, p_2, \Omega_0, t_E)} = 0 \tag{4.23}$$

The above gives the following four simultaneous equations:

$$\begin{aligned}
\overline{\lambda t} - \bar{\lambda} t_E - p_1(\bar{t}^2 - 2t_E \bar{t} + t_E^2) &= 0 \\
\overline{\Omega t} - \bar{\Omega} t_E - p_2(\bar{t}^2 - 2t_E \bar{t} + t_E^2) &= \Omega_0(\bar{t} - t_E) \\
\bar{\Omega} - \Omega_0 - p_2(\bar{t} - t_E) &= 0 \\
p_1 \bar{\lambda} + p_2 \bar{\Omega} &= p_2 \Omega_0 + (p_1^2 + p_2^2)(\bar{t} - t_E)
\end{aligned} \tag{4.24}$$

where, the bar indicates an average over all data points from Equation: 4.18. These equations are solved algebraically to determine p_1, p_2, Ω_0 and t_E :

$$\begin{aligned}
p_1 &= \frac{1}{(\bar{t}^2 - \bar{t}^2)} (\bar{\lambda}t - \bar{\lambda}\bar{t}) \\
p_2 &= \frac{1}{(\bar{t}^2 - \bar{t}^2)} (\bar{\Omega}t - \bar{\Omega}\bar{t}) \\
t_E &= \bar{t} - \frac{\bar{\lambda}}{p_1} \\
\Omega_0 &= \bar{\Omega} - \frac{p_2}{p_1} \bar{\lambda}
\end{aligned} \tag{4.25}$$

4.3.3 Determination of Semi Major Axis “ a ” and Inclination “ I_0 ”

The slope estimates of p_1 and p_2 have an implicit dependence upon a and I_0 as expressed in Equation: 4.2. Therefore, one need to determine a and I_0 to keep the secular terms of κ and ϑ accurate. The estimate of I_0 is found out based on estimate of p_2 (Equation: 4.25) in an iterative scheme wherein this estimate and the analytical expression for the slope “ ϑ ” obtained from an analytical equation of Ω (Equation: 4.13) are equated fixing the semi major axis “ a ”. One starts by assuming a value of “ a ” and it is convenient to choose the first value of the radial coordinate $r(t_0)$ from Equation: 4.18.

$$\therefore \vartheta = \frac{p_2}{n} \equiv \vartheta_2 + (\vartheta - \vartheta_2) \tag{4.26}$$

where, ϑ_2 is the value of ϑ for just J_2 . This is much larger than $(\vartheta - \vartheta_2)$. Rewriting the equation for ϑ_2 as a function of x , from Equation: 4.2 one obtains:

$$\vartheta_2 = -\frac{3}{2}J_2 \left(\frac{R}{a}\right)^2 x \tag{4.27}$$

where, $x = \cos I_0$, one can get first estimate of x from:

$$\vartheta_2(x_1) = \frac{p_2}{n} \tag{4.28}$$

The iterative scheme from Equation: 4.26 for i^{th} estimate of $\vartheta_2(x_i)$ would be:

$$\vartheta_2(x_i) = \frac{p_2}{n} - [\vartheta(x_{i-1}) - \vartheta_2(x_{i-1})] \tag{4.29}$$

The iteration of Equation: 4.29 for ϑ_2 should be carried out until the change in result decreases below a selected threshold. One may now use the value of ϑ_2 to find value of I_0 from Equation: 4.27.

The value of p_1 (Equation: 4.25) would now be used to estimate the semi-major axis a . This can be done by using the Newton-Raphson method [100]. From Equation; 4.21 Let,

$$f = 1 + \kappa - p_1 a^{3/2} \mu^{-1/2} \quad (4.30)$$

When a satisfies the relation $(1 + \kappa)n = p_1$, then $f = 0$. Equation: 4.30 can be written for the Newton-Raphson formulation as:

$$a_i = a_{i-1} - \frac{f(a_{i-1})}{\dot{f}_a(a_{i-1})} \quad (4.31)$$

where, $\dot{f}_a(\cdot)$ indicates the derivative of a function with respect to a . After some derivations of the i^{th} estimate of a may be written as:

$$a_i = a_{i-1} + a_{i-1} \frac{1 + \kappa - p_1 a^{3/2}}{\kappa' + \frac{3}{2} p_1 a^{3/2}} \quad (4.32)$$

where, $\kappa' = \sum_{m=1}^M 2m\kappa_{2m} + 4\kappa_{2s}$, and κ_{2s} is the J_2^2 coefficient in the post epicycle equation (see Appendix-C). Again Equation: 4.32 will be iterated until the change in the semi-major axis “ a ” is less than a selected tolerance. Now with the newly found out value of “ a ” one substitutes this value of “ a ” in Equation: 4.29 for I_0 and repeat this procedure until both the values a and I_0 converge.

4.3.4 Determination of “ ξ ” and “ η_P ”

The quantities of ξ_P and η_P are now being estimated using the equations of the epicycle coordinates of r and v_r . The equation of these coordinates can be expressed as [14]:

$$\begin{aligned} r &= a[(1 + \rho) - \xi_P \cos \alpha - \eta_P \sin \alpha + \chi \sin \beta + \Delta r_2 \cos 2\beta + \Delta_r] \\ v_r &= an[\xi_P \sin \alpha - \eta_P \cos \alpha - (1 + \kappa)(2\Delta_r \sin 2\beta - \chi \cos \beta) + \dot{\Delta}_r] \end{aligned} \quad (4.33)$$

By separating out oscillating terms, these equations may be rewritten as:

$$\begin{aligned} \frac{r}{a} &= dr - \xi_P \cos \alpha - \eta_P \sin \alpha \\ \frac{v_r}{an} &= dv_r + \xi_P \sin \alpha - \eta_P \cos \alpha \end{aligned} \quad (4.34)$$

where,

$$\begin{aligned} dr &= 1 + \rho + \chi \sin \beta + \Delta r_2 \cos 2\beta + \Delta_r \\ dv_r &= -(1 + \kappa)(2\Delta r_2 \sin 2\beta - \chi \cos \beta) + \dot{\Delta}_r \end{aligned} \quad (4.35)$$

The above equations can be further simplified on the lines of Equation: 4.19 to 4.21, as:

$$\begin{aligned} \therefore \frac{r}{a} - dr &\equiv \tilde{r} = -\xi_P \cos \alpha - \eta_P \sin \alpha \\ \frac{\dot{r}}{an} - dv_r &\equiv \tilde{v} = \xi_P \sin \alpha - \eta_P \cos \alpha \end{aligned} \quad (4.36)$$

One may now conveniently define the least square cost function as:

$$\mathfrak{J} = \sum_{i=0}^k [\tilde{r}(t_i) + \xi_P \cos \alpha_i + \eta_P \sin \alpha_i]^2 + \sum_{i=0}^k [\tilde{v}_r(t_i) - \xi_P \sin \alpha_i + \eta_P \cos \alpha_i]^2 \quad (4.37)$$

where, $\alpha_i = n(t_i - t_E)$. The above function should be differentiated with respect to ξ_P and η_P and set equals to zero:

$$\frac{\partial \mathfrak{J}}{\partial (\xi_P, \eta_P)} = 0 \quad (4.38)$$

The resultant simultaneous equation can be solved to provide estimates of ξ_P , η_P expressed as:

$$\begin{aligned} \xi_P &= \overline{v \sin \alpha} - \overline{r \cos \alpha} \\ \eta_P &= -\overline{v \cos \alpha} - \overline{r \sin \alpha} \end{aligned} \quad (4.39)$$

Now these parameters can be conveniently used in Equation: 4.20 to compute $d\lambda$ and $d\Omega$ and the estimates of secular terms would be repeated as in Equations 4.29 and 4.32. The algorithm is repeatedly executed until the estimates are converged to the orbital parameters. The estimated parameters are denoted as $\hat{\mathbf{x}}_0$. See Figure: 4-3 for the flow chart of the EPF.

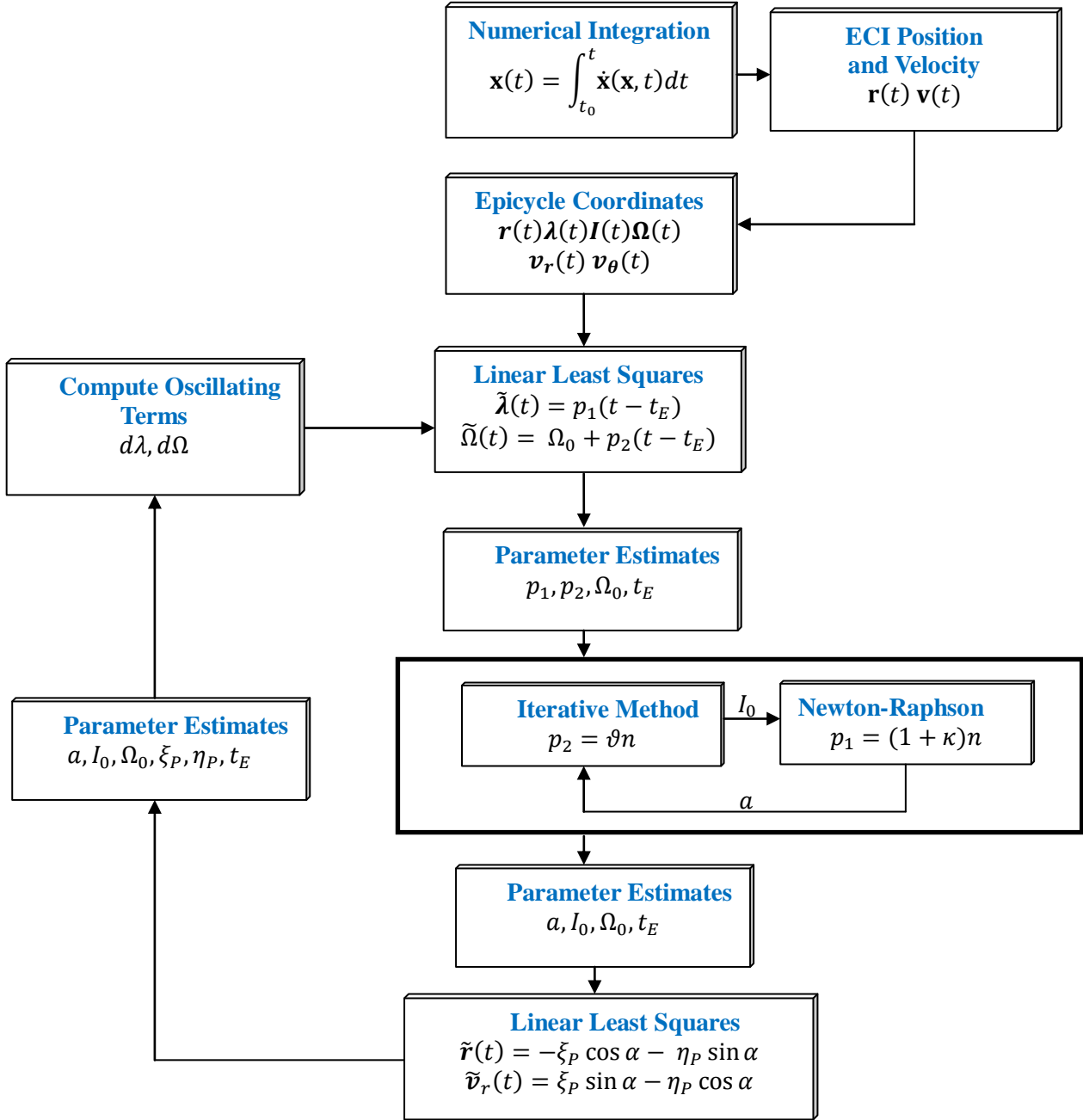


Figure 4-3: Flow chart of the Epicycle Parameter Filter (EPF). Important features of the algorithm includes the use of numerically obtained epicycle orbital coordinates in linear least squares formulation to compute p_1 and p_2 (slopes of linear growth in coordinates of λ and Ω), compute semi-major axis “ a ” and inclination “ I_0 ” by using iterative methods and Newton-Raphson root finding algorithm and linear least squares to compute ξ_P and η_P .

4.4 Parameter Estimation Accuracy

The fundamental idea of EPF is to accurately estimate the coefficients for secular variations i.e., κ and ϑ . Since, these two coefficients are functions of “ a ” and “ I_0 ”, therefore, it seems appropriate to validate estimation accuracies by varying these two parameters. A useful insight on variation and strength of these coefficients can be obtained by plotting them as a function of I_0 and a (see Figure: 4-4 and 4-5).

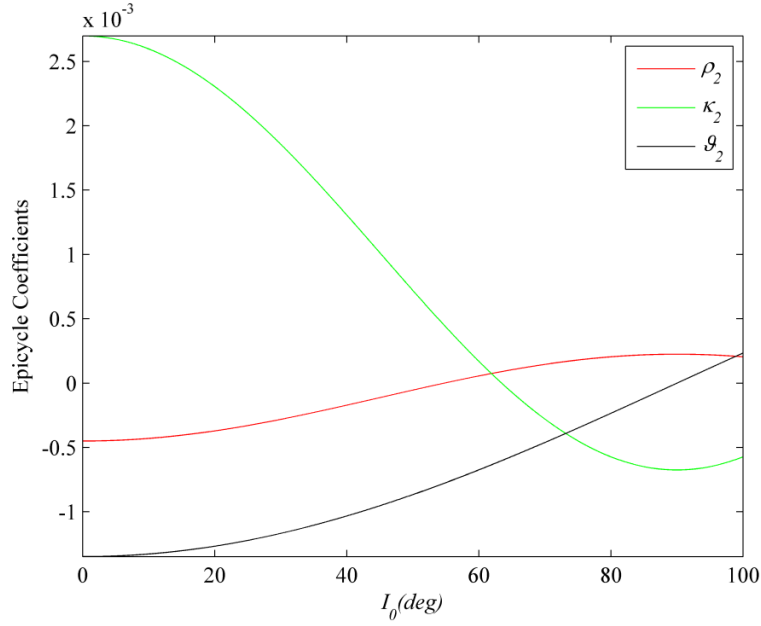


Figure 4-4: J_2 epicycle coefficients for radial offset (ρ_2), and secular drift (κ_2, ϑ_2), are plotted as a function of the inclination I_0 by fixing $a = 7000$ km.

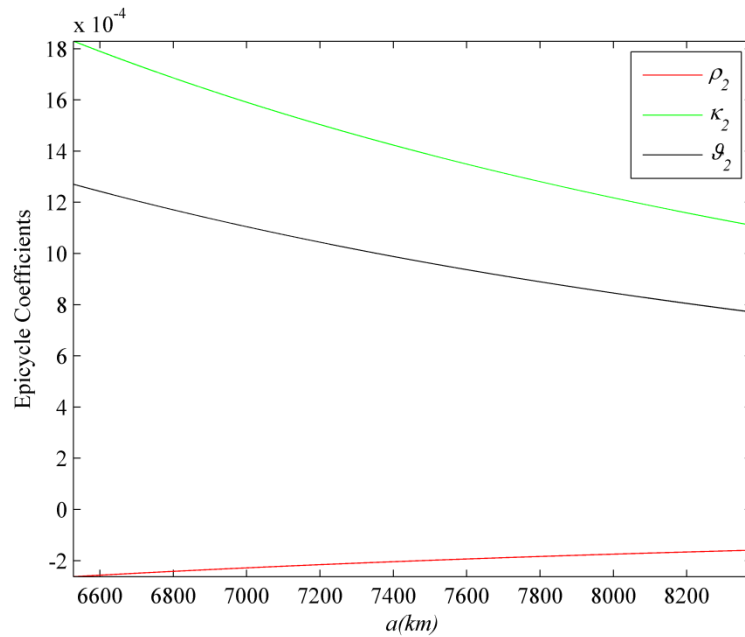


Figure 4-5: J_2 epicycle coefficients for the radial offset (ρ_2), and secular drift (κ_2, ϑ_2), are plotted as a function of “ a ” by fixing $I_0 = 98$ deg.

The most significant perturbative term J_2 has been used to compute epicycle coefficients in Figures: 4-4 and 4-5. One may clearly observe higher amplitudes of these coefficients at lower inclinations and lower semi-major axis (LEO). Moreover, changes in I_0 (Figure: 4-4) suggests more rapid and significant, secular effects on satellite orbits compared with changes in the semi-major axis “ a ”. Therefore, one would consider different choices of I_0 for all inclinations and observe the estimation accuracies.

The semi-major axis is chosen as 7000 km orbit and one considers the zonal harmonic terms up to J_4 of non-spherical geopotential for the experiments [38]. This would serve as a sequel to the generic scheme for estimation, developed earlier in the chapter for higher order harmonics. By taking the gradient of gravitational potential function “ \mathcal{U} ” expressed in Equation: 3.31, one may express the \mathbf{a}_G term in Equation: 4.15 for order up to J_4 of the geopotential as under [12].

$$J_2(\mathbf{r}) = -\frac{3J_2\mu_E R_E^2}{2r^5} \begin{Bmatrix} \left(1 - \frac{5Z^2}{r^2}\right)X \\ \left(1 - \frac{5Z^2}{r^2}\right)Y \\ \left(3 - \frac{5Z^2}{r^2}\right)Z \end{Bmatrix} \quad (4.40)$$

$$J_3(\mathbf{r}) = -\frac{5J_3\mu_E R_E^3}{2r^7} \begin{Bmatrix} \left(3Z - \frac{7Z^3}{r^2}\right)X \\ \left(3Z - \frac{7Z^3}{r^2}\right)Y \\ \left(6Z^2 - \frac{7Z^4}{r^2} - \frac{3}{5}r^2\right) \end{Bmatrix} \quad (4.41)$$

$$J_4(\mathbf{r}) = \frac{15J_4\mu_E R_E^4}{8r^7} \begin{Bmatrix} \left(1 - \frac{14Z^2}{r^2} + \frac{21Z^4}{r^4}\right)X \\ \left(1 - \frac{14Z^2}{r^2} + \frac{21Z^4}{r^4}\right)Y \\ \left(5 - \frac{70Z^2}{3r^2} + \frac{21Z^4}{r^4}\right)Z \end{Bmatrix} \quad (4.42)$$

where X , Y , and Z are the coordinates in ECI frame (details in Chapter: 3), $r = \sqrt{X^2 + Y^2 + Z^2}$, $J_2 = 1.08262668355 \times 10^{-3}$, $J_3 = -2.53265648533 \times 10^{-6}$, and $J_4 = -1.61962159137 \times 10^{-6}$.

The estimation of the parameters ($\hat{\mathbf{x}}_0$) is compared with the true parameters (\mathbf{x}_0) used for numerical propagation of the reference trajectory (Equation: 4-18). The batch of numerically propagated data used for estimation is for a week. The time scale of 1 week corresponds to approximately 100 orbital periods for LEO micro-satellites (weighs between 10 to 100 kg) or nano-satellites (weighs less than 10 kg) [102]. Figure: 4-6 indicates percentage errors in estimating the parameters of a , Ω_0 , and α_0 ; whereas, Figure: 4-7 illustrates errors for I_0 , ζ_P , and η_P .

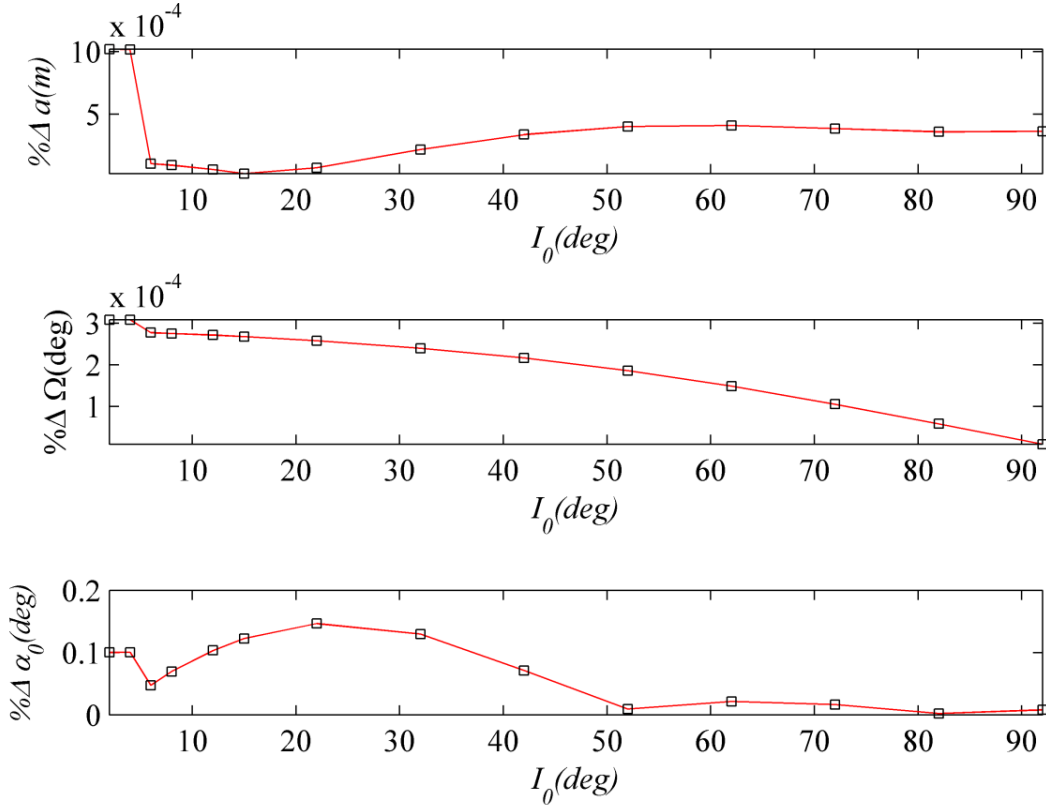


Figure 4-6: Percentage estimation errors (Δ) for semi-major axis " a "(top), right ascension of the ascending node " Ω_0 "(middle), and initial epicycle phase " α_0 "(bottom), as a function of inclination of the orbital plane.

Results reveal that errors cannot be fully eliminated as the two trajectories are being propagated differently i.e., in the numerical and analytical solutions. In order to keep the two trajectories sufficiently close to each other for a long duration slightly perturbed parameters are found out in order to compensate for the process noise [5]. In general the, process noise is a time varying quantity and is

inherent to all analytically derived models due to linearization and approximations to full dynamics of the nonlinear problem.

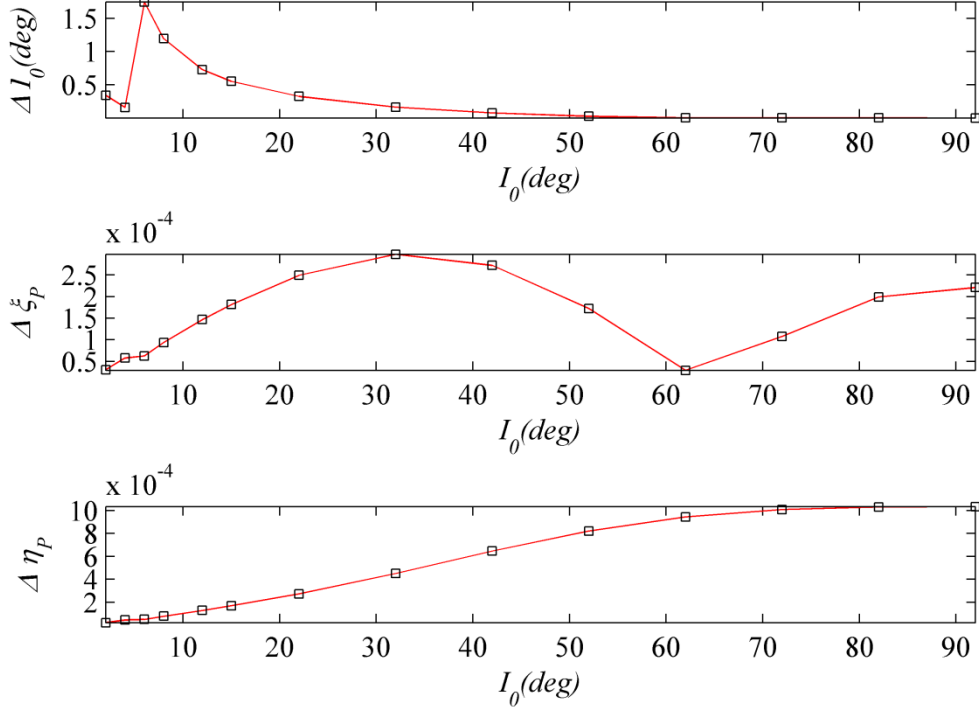


Figure 4-7: Estimation errors (Δ) for inclination " I_0 " (top), " ξ_P " (middle) and " η_P " (bottom) as a function of inclination of the orbital plane.

4.5 Error Statistics in Orbital Coordinates at Different I_0

The optimal parameter estimates ($\hat{\mathbf{x}}_0$) for all orbital inclinations, as discussed in the previous section are now being used for generation of analytic epicycle trajectory using the evolution equations (Equation: 3.35). The analytic epicycle coordinates are expressed as:

$$\begin{aligned} \mathcal{P}_{Epicycle}(\hat{\mathbf{x}}_0, t) = & [(r^p(t_0), \lambda^p(t_0), I^p(t_0), \Omega^p(t_0), v_r^p(t_0), v_\theta^p(t_0)) \dots \\ & \dots (r^p(t_k), \lambda^p(t_k), I^p(t_k), \Omega^p(t_k), v_r^p(t_k), v_\theta^p(t_k))]^T \end{aligned} \quad (4.43)$$

where, $\mathcal{P}_{Epicycle}(\hat{\mathbf{x}}_0, t)$ is a vector containing the coordinates of the analytic epicycle trajectory, and superscript "p" stands for analytic epicycle coordinates.

The numerical trajectory available from Equation: 4.18 would be used to compute the error statistics

in terms of epicycle coordinates by subtracting numerical coordinates and analytic coordinates from Equation: 4.43. The maximum errors for the period of one week are observed only at lower inclinations. The errors decrease almost exponentially at higher inclinations. Figures: 4-8 to 4-10, illustrate the maximum errors in epicycle coordinates over the period of one week.

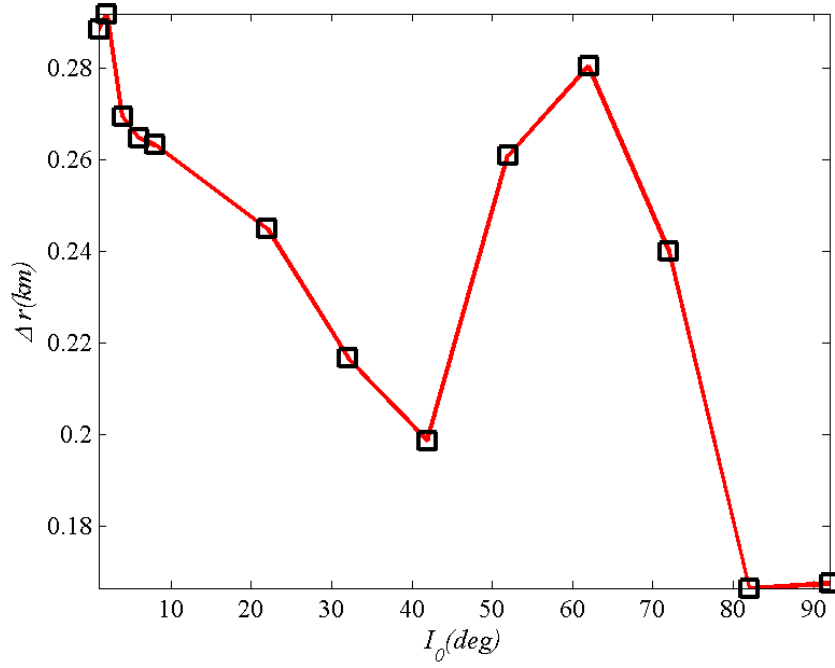


Figure 4-8: Maximum absolute errors in “ r ” as function of inclination of the orbital plane.

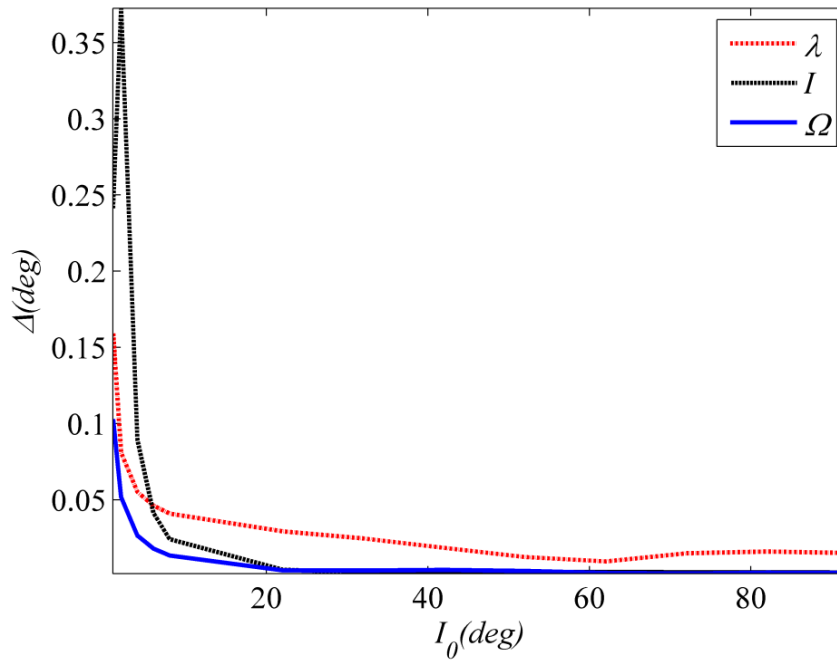


Figure 4-9: Maximum absolute errors in λ , I , and Ω as a function of inclination of the orbital plane.

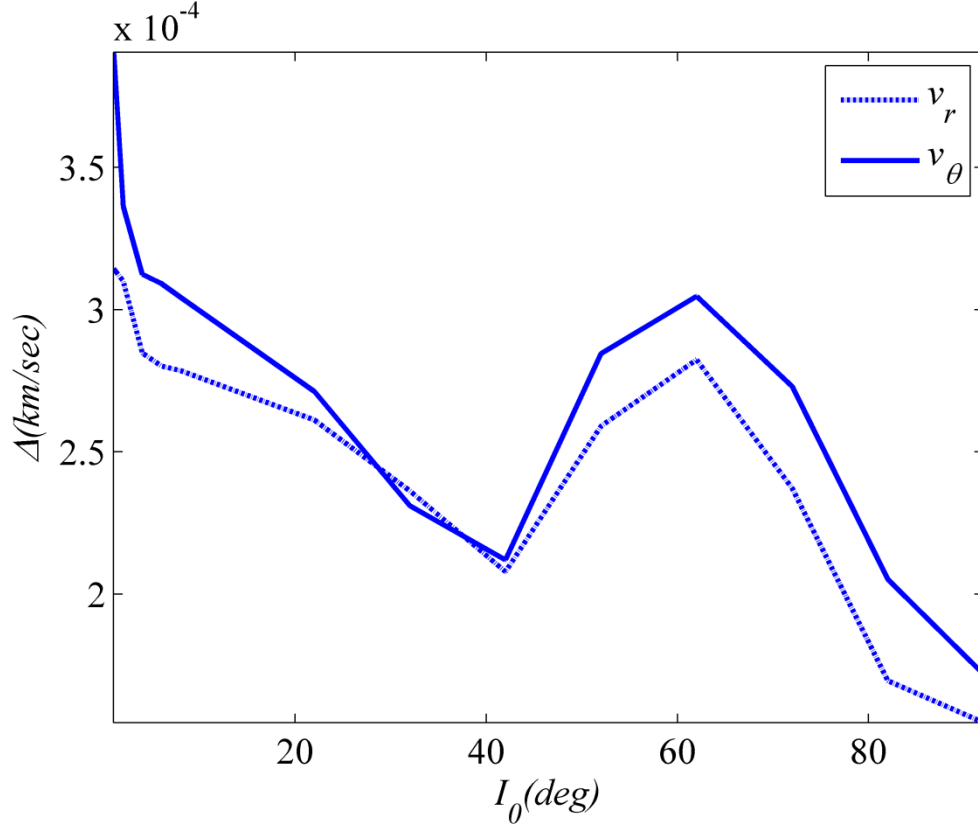


Figure 4-10: Maximum absolute errors in v_r and v_θ as a function of inclination of the orbital plane.

The errors in radial coordinate (Δr) (Figure: 4-8) and radial / azimuthal velocity (v_r, v_θ) (Figure: 4-10) shows an unexpected increase in errors in the vicinity of critical inclination ($I_0 = 63.4$ deg). This is possibly due to the approximation of long periodic variations for J_3 given by [2]:

$$\chi_3 = \left(\frac{J_3}{2J_2} \right) \left(\frac{R_E}{a} \right) \sin I_0 \quad (4.44)$$

The approximation is carried out in order to avoid the term κ_2 in the denominator of Equation: 4.11, getting zero at $\sin^2 I_0 = 4/5$ (see Appendix-C for expression of κ_2). This happens to be at $I_0 = 63.43$ deg. Nevertheless the errors are small and can be minimized by replacing $(1/\kappa_2)$ with $(1/\kappa)$. Thereby, including higher order harmonics one can avoid such a numerical instability.

4.6 Time History of Errors in Epicycle Coordinates

In order to observe the time history of errors in epicycle coordinates a sun synchronous LEO satellite

with $a = 7003 \text{ km}$, and following initial conditions in ECI coordinates $(\mathbf{r}_0, \mathbf{v}_0)^T$ is selected:

$$\mathbf{r}_0 = [6898.3267 \quad 71.6119 \quad 1203.7126]^T \text{km} \quad (4-45)$$

$$\mathbf{v}_0 = [-1.2740 \quad -1.0793 \quad 7.3589]^T \text{km/s}$$

The time history of errors for the period is shown in Figures: 4-11 to 4-16. The estimates are approximately zero mean and converged which shows consistency in estimates. The maximum absolute errors over this period are $\Delta r = 183 \text{ m}$, $\Delta \lambda = 15.9 \times 10^{-3} \text{ deg}$, $\Delta I = 2.4 \times 10^{-3} \text{ deg}$, $\Delta \Omega = 2.0 \times 10^{-3} \text{ deg}$, $\Delta v_r = 2.1 \times 10^{-4} \text{ km/s}$, $\Delta v_\theta = 2.0 \times 10^{-4} \text{ km/s}$.

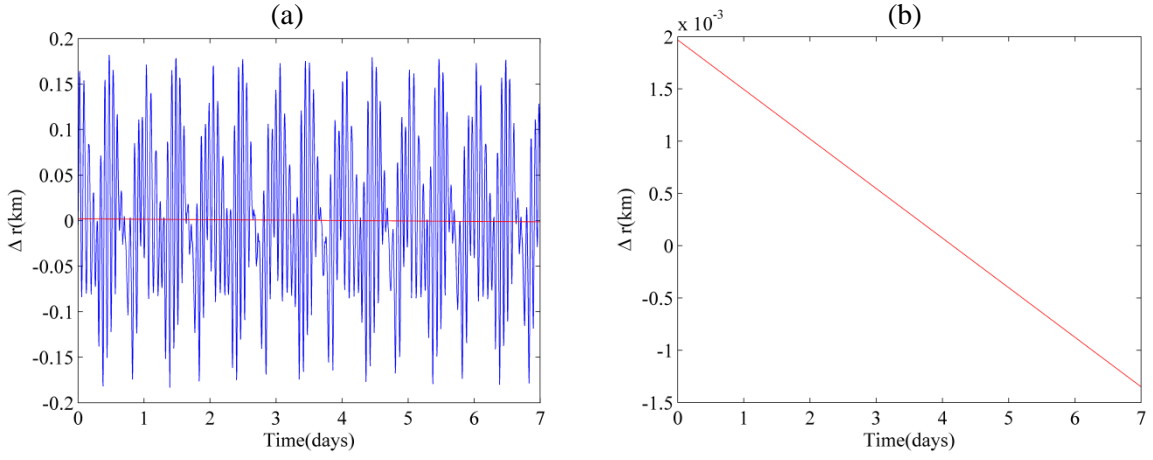


Figure 4-11: (a) Time history of errors (Δ) in “ r ” using optimal $\hat{\mathbf{x}}_0$. Note convergence and negligible drift in mean error in epicycle radial coordinate (shown as red line). (b) Mean error is identical around zero mean value over the simulation time of one week.

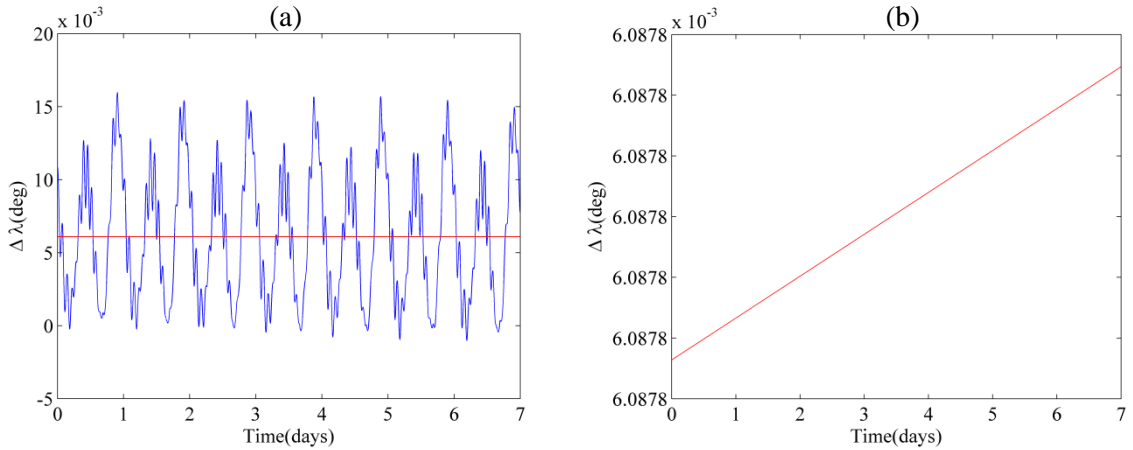


Figure 4-12: (a) Time history of errors (Δ) in argument of latitude λ for a period one week using optimal $\hat{\mathbf{x}}_0$. Note Convergence and constant offset order $\approx 10^{-3} \text{ deg}$ in mean error. (b) The drift in mean error is $< 10^{-3} \text{ deg}$.

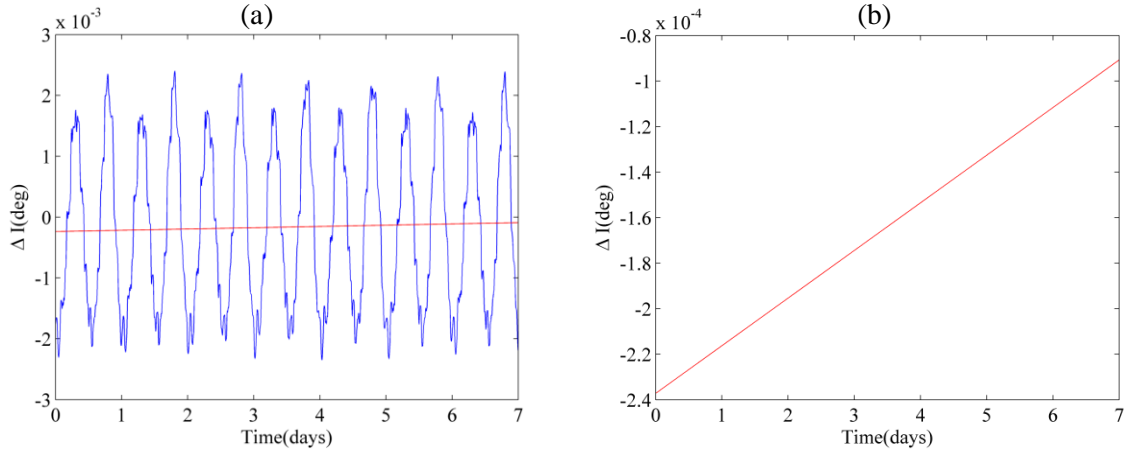


Figure 4-13: (a) Time history of errors (Δ) in inclination “ I ” for a period of one week using optimal $\hat{\mathbf{x}}_0$. Note convergence pattern and mean error offset order $\approx 10^{-4}$ deg. (b) Drift in mean error is $\approx 1.6 \times 10^{-4}$ deg.

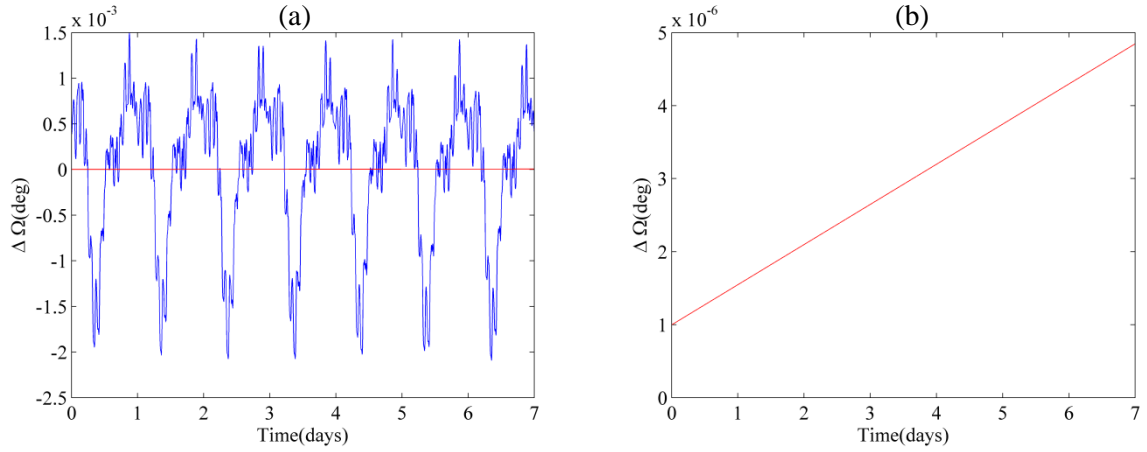


Figure 4-14: Time history of errors (Δ) in RAAN “ Ω ” for a period of one week using optimal $\hat{\mathbf{x}}_0$. Note convergence and mean error offset order $\approx 10^{-6}$ deg. (b) Drift in mean error is $\approx 4 \times 10^{-6}$ deg.

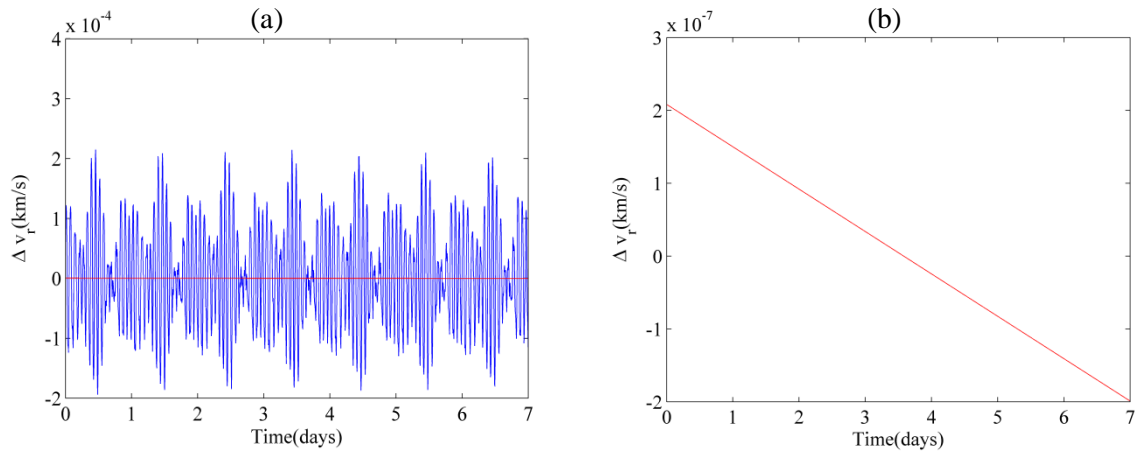


Figure 4-15: Time history of errors (Δ) in v_r for a period of one week using optimal $\hat{\mathbf{x}}_0$. Note convergence and mean error offset order $\approx 10^{-7}$ km/s. (b) Drift in mean error is only $\approx 4 \times 10^{-7}$ km/s.

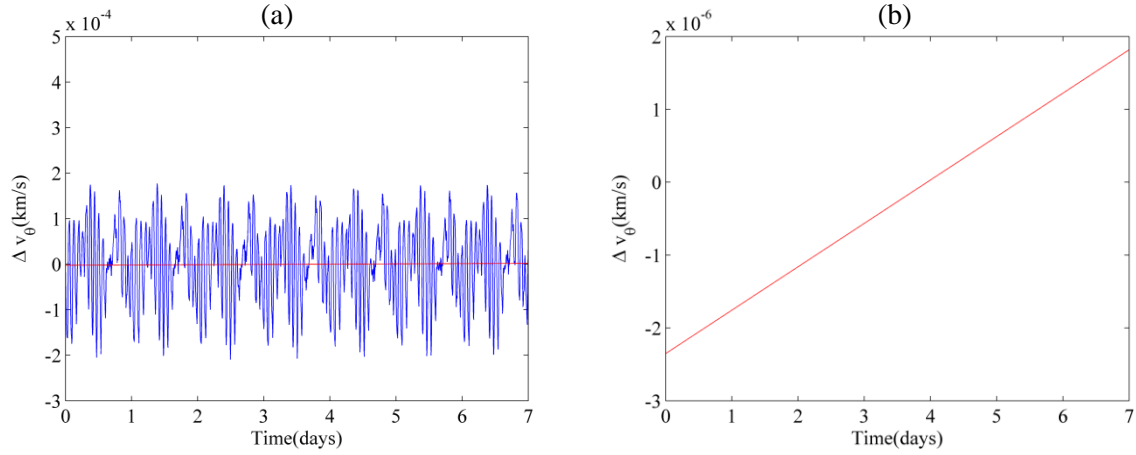


Figure 4-16: Time history of errors (Δ) in v_θ for a period of one week using optimal $\hat{\mathbf{x}}_0$. Note convergence and mean error offset order $\approx 10^{-6}$ km/s. (b) Drift in mean error is $\approx 4 \times 10^{-6}$ km/s.

4.7 Time History of Errors in Epicycle Coordinates Without Estimation

The time history of errors will now be observed for initial conditions as expressed in Equation: 4-45 without using EPF. The errors in the coordinates are presented in Figures: 4-17 to 4-22. One may clearly observe the increase and drift in errors for all the coordinates if the orbital parameters are not properly selected. Divergence and increased errors are quite evident from these plots; especially in the argument of latitude and the right ascension of the ascending node which amounts to significant in-track and cross-track errors in LVLH coordinate system.

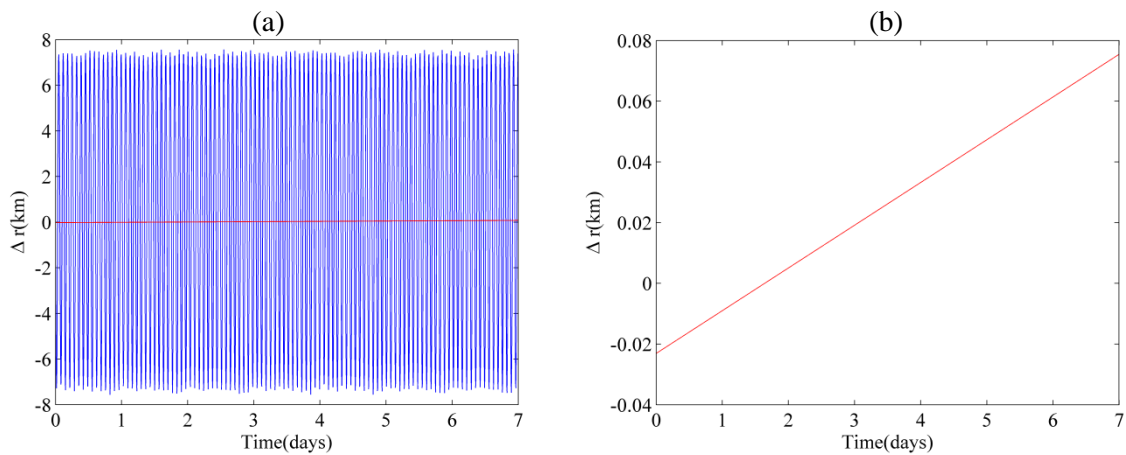


Figure 4-17: Time history of errors (Δ) in “ r ” without estimation. Notice increased divergence of mean error (shown as red line) and error oscillations ≈ 8 km once compared with Figure: 4-11.

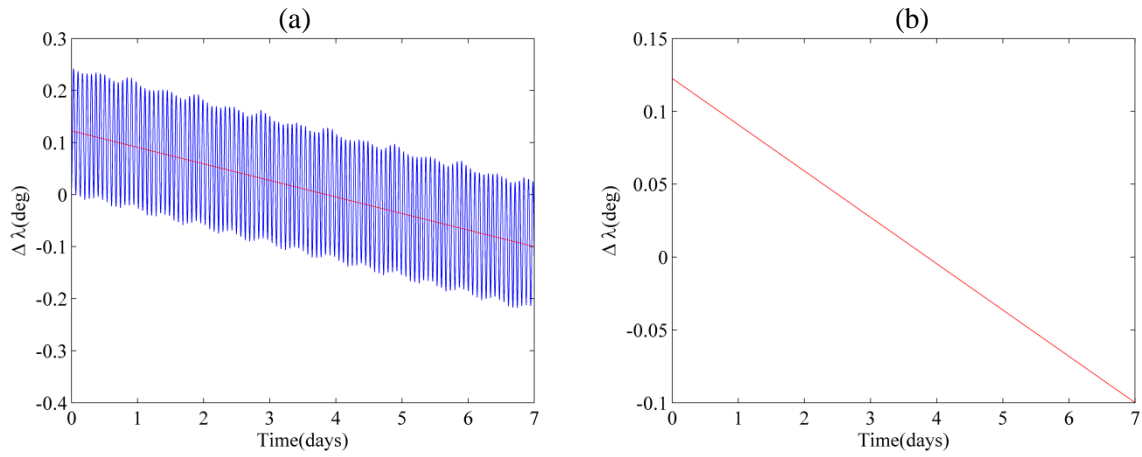


Figure 4-18: Time history of errors (Δ) in “ λ ” without estimation. Notice the significant divergence and periodic errors once compared with Figure: 4-12.

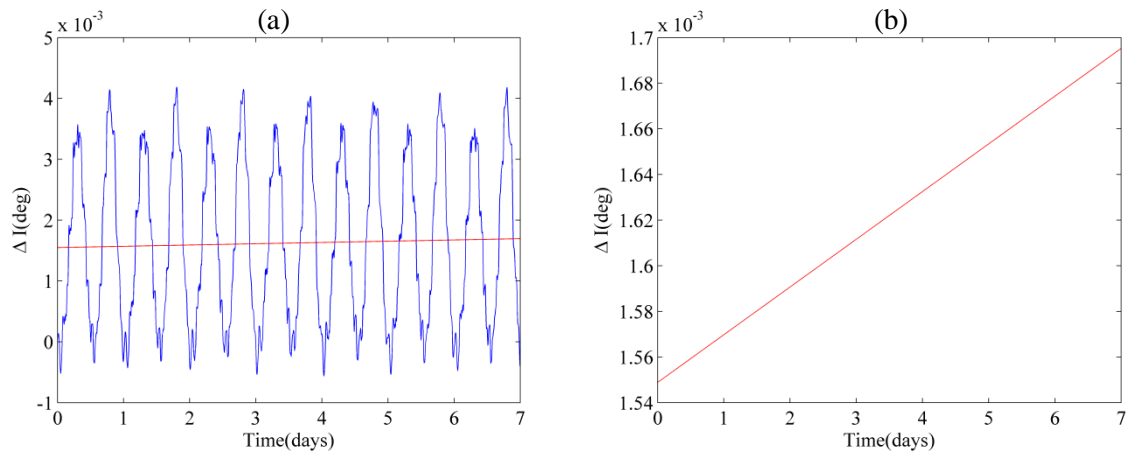


Figure 4-19: Time history of errors (Δ) in “ P ” without estimation. Notice the increased error oscillations and drift in mean error compared with Figure: 4-13.

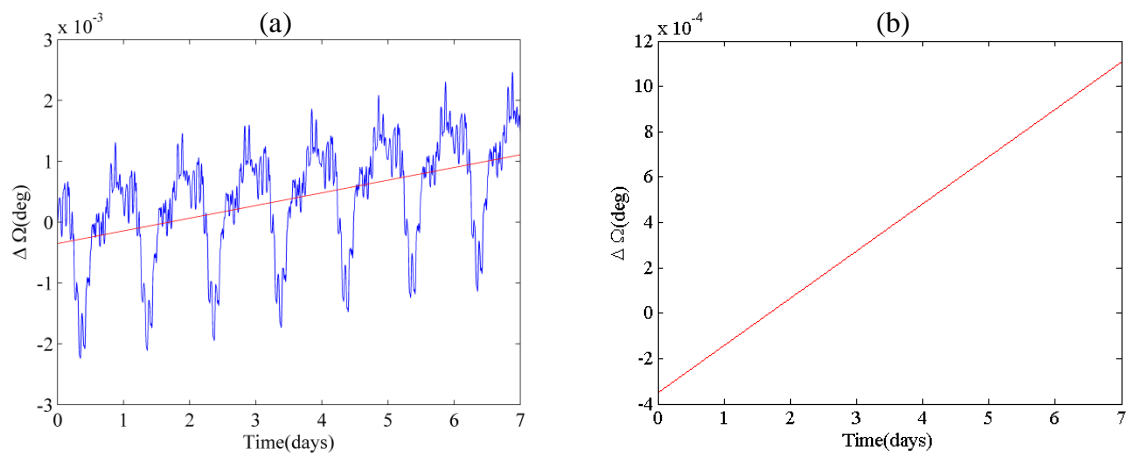


Figure 4-20: Time history of errors (Δ) in “ Ω ” without estimation. Notice the divergence and increased periodic errors once compared with Figure: 4-14.

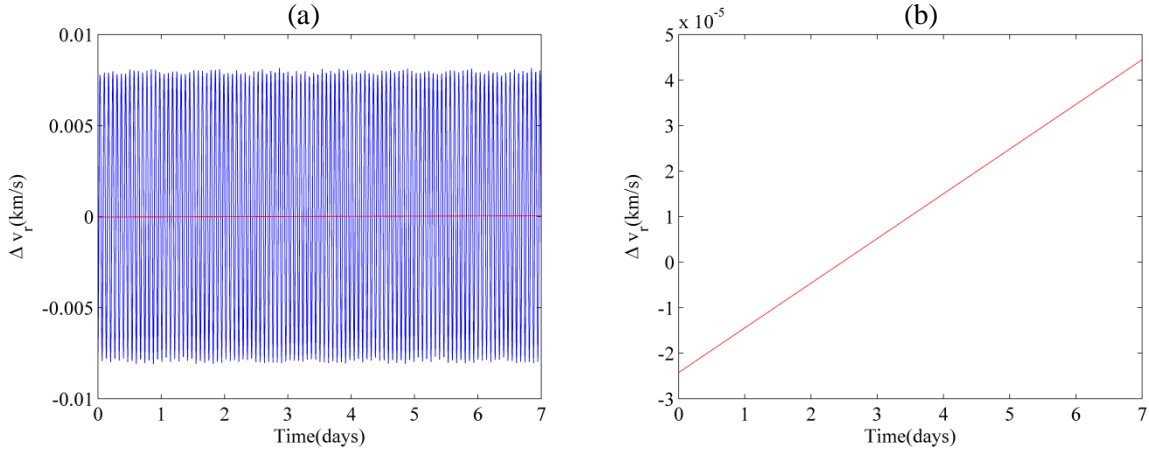


Figure 4-21: Time history of errors (Δ) in v_r without estimation. Notice the increased error oscillations and drift once compared with Figure: 4-15.

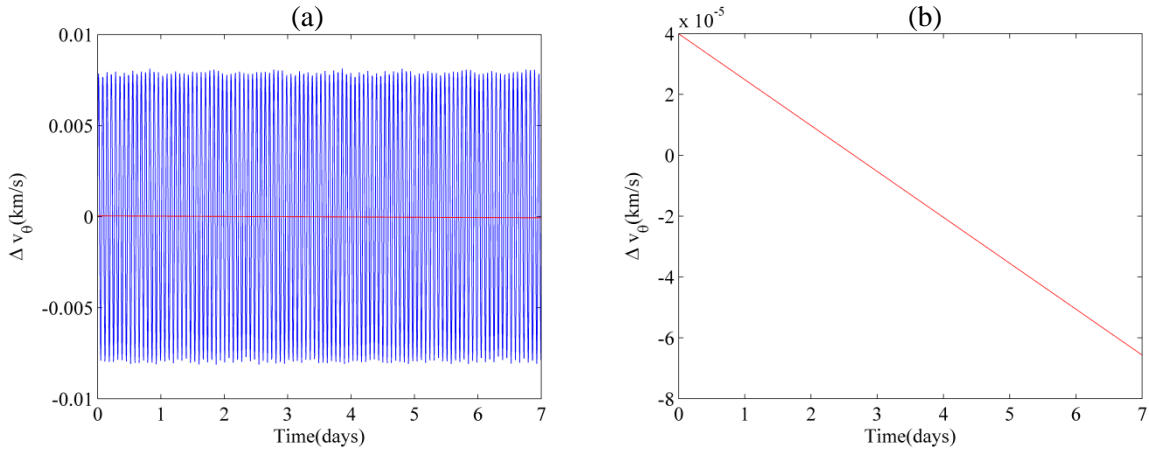


Figure 4-22: Time history of errors (Δ) in v_θ without estimation. Notice the increased error oscillations and drift once compared with 4-16.

4.8 Free Propagation Secular Error Growth

The forward evolution of epicycle position coordinates (r, λ, I, Ω) will now be observed for the growth of errors after having been initialized with the optimal parameters using an EPF. The choice of such optimal parameters is based on the orbital data of one week. The *error growth* criterion is selected as drift (secular growth) in mean errors by 10% of the maximum error in a particular position coordinate. Essentially one would observe the time by which the drift in mean errors in a particular position coordinate exceeds the *error growth* criterion. This would form another useful measure of efficiency for the linear filter. The drift in mean errors for each position coordinate is computed using a linear least squares approximation [99] (see Figures: 4-23 to 4-26). See Table 4-1 for the *error growth* criterion at the end of 12th day. This means one would have to re-estimate the epicycle parameters at the end of 12th day as the growth in λ exceeds 10% at that time.

Position Coordinate	Drift of error (percentage of the maximum error)
r	1.22%
λ	10.81%
I	3.11%
Ω	6.1%

Table 4-1: The table shows the drift of error in terms of percentage of the maximum error in a particular coordinate at the end of 12th day.

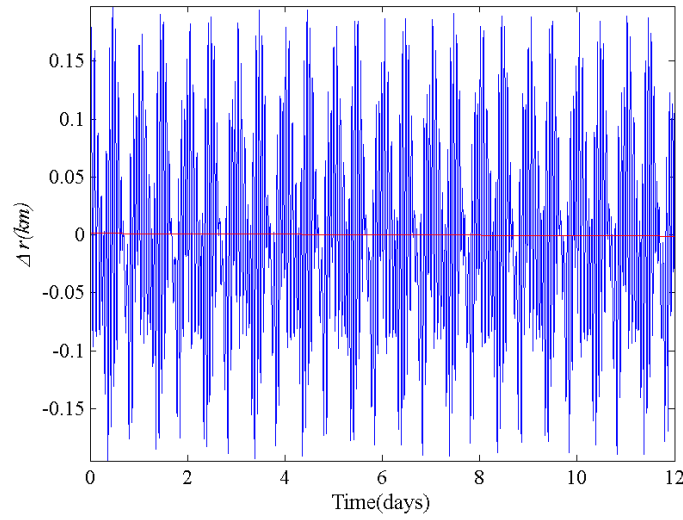


Figure 4-23: Time history of radial coordinate error (Δ) over 12 days. The red line shows linear growth / drift computed using least squares approximation. The drift is about 1.22% of the maximum error at the end of 12th day.

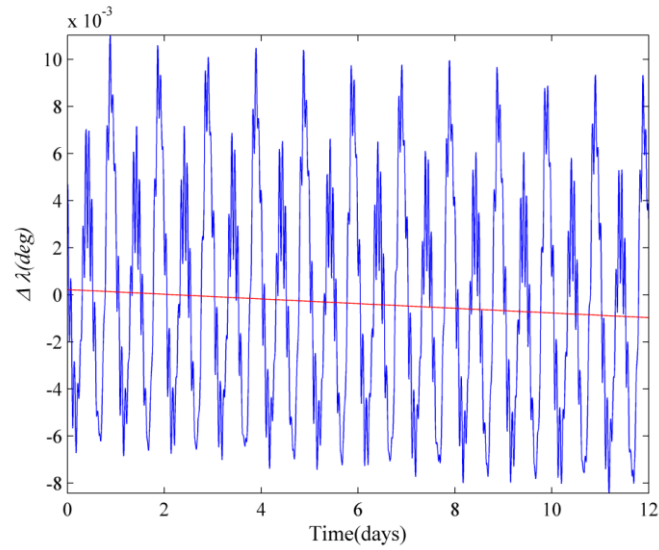


Figure 4-24: Time history of errors (Δ) in argument of latitude over 12 days. The red line shows drift in errors, computed using linear least squares approximation. The drift is about 10.81% of the maximum error at the end of 12th day.

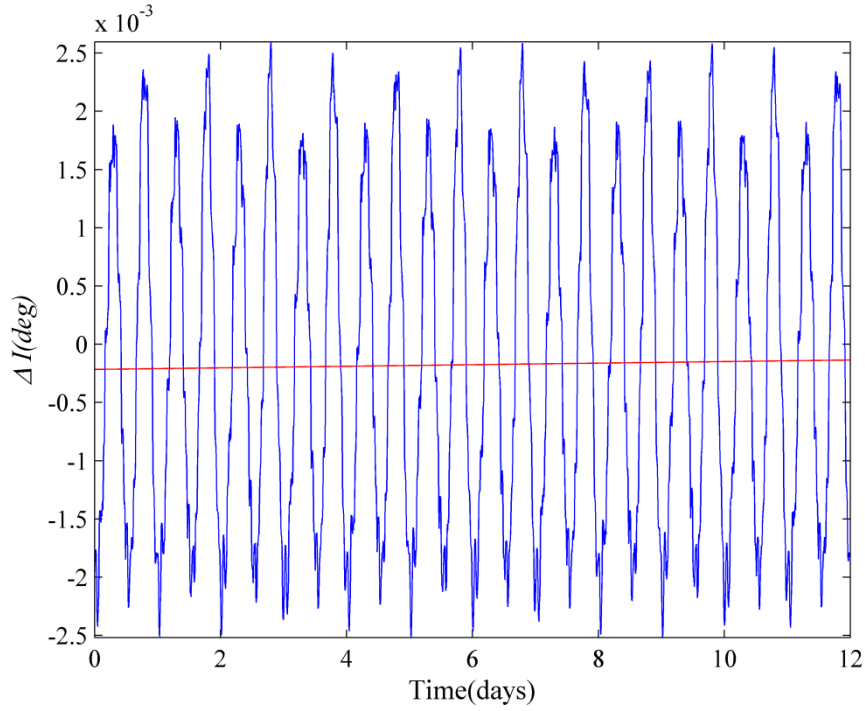


Figure 4-25: Time history of errors (Δ) in inclination over 12 days. The red line shows drift in errors, computed using linear least squares approximation. The drift is about 3.11% of the maximum error at the end of 12th day.

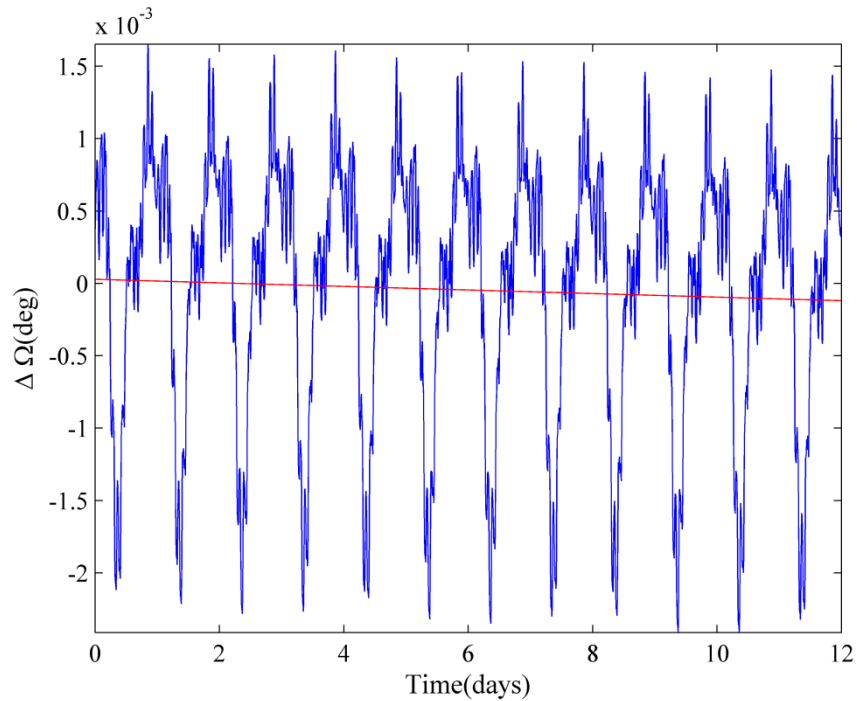


Figure 4-26: Time history of errors (Δ) in RAAN over 12 days. The red line shows drift in errors, computed using linear least squares approximation. The drift is about 6.1% of the maximum error at the end of 12th day.

4.9 Summary

This chapter discusses the development of an EPF for epicycle orbits including higher order zonal harmonic terms. As an example the methodology of EPF has been simulated for terms up to J_4 . Nevertheless, the higher order perturbative terms can easily be extended using Equations: 4.1 and 4.6 for use by EPF. The estimation results show improved epicycle coordinates compared to the nonlinear numerical trajectory. The maximum errors were reduced as 97% in r , 93% in λ , 41% in I , 16% in Ω , 97% in v_r and 97% in v_θ . By keeping the drift in the mean errors as 10% of the maximum error in a particular position coordinate, repeated estimation of the epicycle parameters would be needed after *twelve days*. The repeated parameter estimates can be performed on ground stations for later update to satellite onboard Attitude and Orbit Control Systems (AOCS) using telemetry and telecommand communication links. The epicyclic orbit equations (Equation: 3-35) can be used on board as a replacement of high precision computationally expensive numerical propagators. It can be conveniently used for computing epicycle orbital parameters from NORAD TLE fit for long durations [101]. The parameters can be used to update orbital parameters for the space catalogues of commercial and non – commercial spacecrafts. Design constellations based on orbital parameters which are more intuitive rather than using differential equations.

5 Development of Gram Charlier Series and its Mixture Particle Filters

5.1 Introduction

Nonlinear Bayesian state estimation for Discrete State Space Model (DSSM) [22], known as *discrete time* filtering has been briefly introduced in Chapters: 1 and 2. In this chapter, a detailed description on such methods will be carried out with a view to developing SMC [47] estimation algorithms based on GCS [29] and its mixture model [34] approximation of Bayes' *a posteriori* PDF [68]. Continuing the work from Chapters: 3 and 4, where the estimation process is carried out over a batch of nonlinear data, the sequential state estimation is based on processing of an individual nonlinear update, as soon as data is made available. First a review of the fundamentals of *discrete time* filtering and SMC methods will be given. A brief description of the seminal work by reference [41] on SIS-R commonly known as *bootstrap* PF follows and its extension based on Gaussian or Gaussian Mixture Model (GMM) particle filtering [42],[43]. The latter two algorithms based on Gaussian or GMM approximation of Bayes' *a posteriori* PDF can be termed here as *parametric bootstrap* particle filters. Subsequently, this chapter develops new nonlinear Bayesian SMC estimation methods based on GCS and its mixture models. This would form as unification of ideas for improved *parametric bootstrap* particle filtering within the broader context of SMC estimation.

A nonlinear dynamical and measurement system can be formulated as DSSM, expressed as [5],[8],[103],[21]:

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{\Gamma}_k \mathbf{w}_k \quad (5.1)$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \quad (5.2)$$

where, $\mathbf{x}_k \in \mathbb{R}^d$ is the d -dimensional state vector to be estimated, denoted with discrete time subscript " k ", $\mathbf{f}(\cdot) \in \mathbb{R}^{d \times 1}$ is a nonlinear function which evolves the state from $(k^{th} - 1) \rightarrow k^{th}$ discrete instant of time, $\mathbf{\Gamma}_k \in \mathbb{R}^{d \times m}$ is a dispersion matrix, $\mathbf{y}_k \in \mathbb{R}^q$ is a q -dimensional measurement vector, $\mathbf{h}(\cdot) \in \mathbb{R}^{q \times 1}$ is nonlinear measurement function of evolved state, $\mathbf{w}_k \in \mathbb{R}^m$ and $\mathbf{v}_k \in \mathbb{R}^q$ is the m -dimensional and q -dimensional mutually independent additive white Gaussian process and

measurement noise variables, respectively. The whiteness of noise variables is equivalent to requiring the state and measurement sequences to be Markov processes [5] (the development of filtering algorithms is restricted here to such processes only). The state variable \mathbf{x}_k is usually considered as hidden variable, being measured only through \mathbf{y}_k at discrete time instants (see Figure: 5.1).

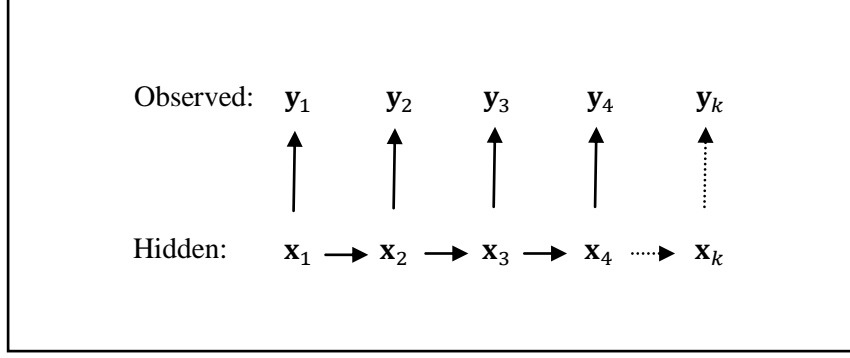


Figure 5-1: In discrete filtering discrete hidden sequence of state \mathbf{x}_k is observed by noisy sequence of observations \mathbf{y}_k . The evolution of state and measurements are obtained at discrete instants of time using positive integer subscripts (1 ... k).

The estimation problem is termed nonlinear if at least one of the models (Equations: 5.1 and 5.2) is the nonlinear function of the state [5]. In a Bayesian framework *a posteriori* PDF of the state $p(\mathbf{x}_k | \mathbf{y}_{1:k})$ given all the observations $\mathbf{y}_{1:k} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$ constitutes the complete solution to the probabilistic inference problem and allows to compute any function of the state $g(\mathbf{x}_k)$ [22]. For example, an optimal estimate of the state $g(\mathbf{x}_k) = \mathbf{x}_k$, in terms of Minimum Mean Square Error (MMSE) estimation criterion would be [5][22]:

$$E[g(\mathbf{x}_k)] \triangleq \hat{\mathbf{x}}_k = \int_{-\infty}^{+\infty} \mathbf{x}_k p(\mathbf{x}_k | \mathbf{y}_{1:k}) d\mathbf{x}_k \quad (5.3)$$

where, $E[.]$ is the expectation operator [99].

The integration of Equation: 5.3 would provide the mean ($\hat{\mathbf{x}}_k$) of Bayes' *a posteriori* PDF termed as MMSE state estimate. The sequential method to obtain Bayes' *a posteriori* PDF as new measurements arrive is achieved by Bayesian recursive formula. By employing Bayes' rule and DSSM as given in Equations: 5.1-5.2 one arrives at following recursive form of *a posteriori* PDF [8],[68],[22]:

$$p(\mathbf{x}_k | \mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{y}_{1:k-1})}{\int_{-\infty}^{+\infty} p(\mathbf{y}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) d\mathbf{x}_k} \quad (5.4)$$

The numerator on right hand side of Equation: 5.4, consists of the likelihood of measurement conditioned on the evolved state $p(\mathbf{y}_k | \mathbf{x}_k)$, and state predictive PDF $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$. The state predictive PDF is obtained through the use of the CKE [5] (Equation: 5.5), using Bayes' *a posteriori* PDF at time instant $k - 1$, expressed as $p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1})$ and the state transition PDF $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ obtained through the nonlinear process model (Equation: 5.1) [22]:

$$p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) = \int_{-\infty}^{+\infty} p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \quad (5.5)$$

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}) = \int_{-\infty}^{+\infty} \delta(\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{w}_k)) p(\mathbf{w}_k) d\mathbf{w}_k \quad (5.6)$$

The likelihood of the measurement conditioned on the evolved state is given by [22]:

$$p(\mathbf{y}_k | \mathbf{x}_k) = \int_{-\infty}^{+\infty} \delta(\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k, \mathbf{v}_k)) p(\mathbf{v}_k) d\mathbf{v}_k \quad (5.7)$$

where, $\delta(\cdot)$ is the Dirac-delta function [104]. Figure: 5-2 depicts the block description of the classic Bayesian recursive filtering methodology.

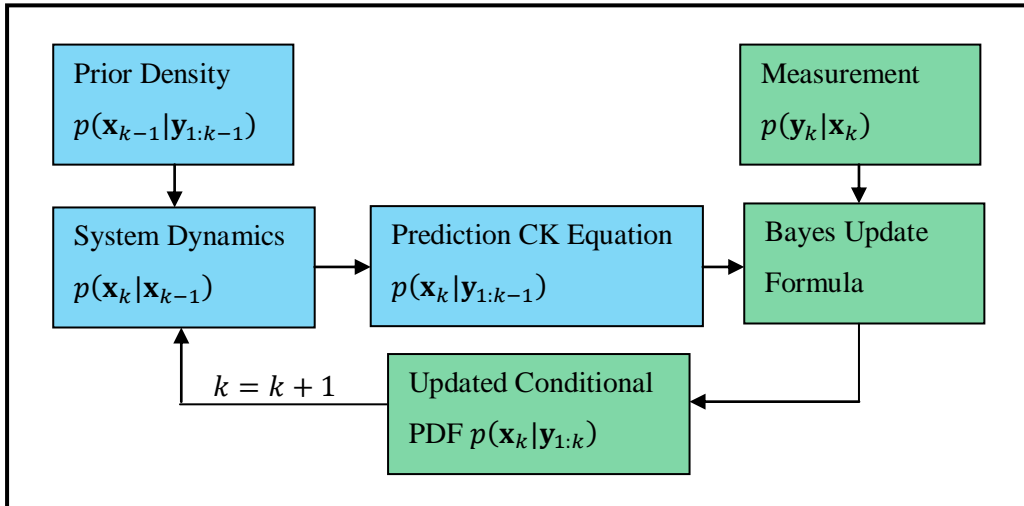


Figure 5-2: Block description of Bayesian prediction and update stages (see text for details).

Equations: 5.4 to 5.7 provide the complete information about the state of a dynamic system in probabilistic sense, from which any type of state inference such as MMSE (Equation: 5.3) or MAP

estimates can be obtained. In this thesis, one would only consider the former type for the development of nonlinear Bayesian filters. The multidimensional integrals in these equations are only tractable for linear dynamic and measurement systems for which the KF is the optimal solution [10]. The KF provides finite dimensional sufficient statistics, comprising of the conditional mean and covariance of the state which completely summarizes the past in a probabilistic sense [5].

Most of the dynamical systems in the real world are nonlinear such as satellite orbital dynamics discussed in Chapter: 3 and 4. If the dynamic system expressed in first of Equation: 5.1 is nonlinear and our belief about its initial conditions and noise distribution are Gaussian or even non-Gaussian, then in general there is no sufficient statistics and Bayesian recursion (Equation: 5.3 to 5.7) has to be used to obtain optimal MMSE estimates. This amounts to an infinite dimensional process in terms of need for an infinite order moment evolution or the requirement to store the entire PDF, which is practically not viable. Therefore, one has to approximate PDFs used in Equations: 5.3 to 5.7 by some tractable form to facilitate solutions for this problem and avoid such a formidable complexity.

Recently a new class of Bayesian filtering methods based on the SMC approach have been considered in the literature called particle filters [41],[68],[42],[22]. SMC methods can be approximately defined as a collection of methods that employs Monte Carlo (MC) simulation scheme in order to fulfil online estimation and prediction requirements [8]. The SMC technique achieves filtering by recursively producing an ensemble of weighted samples termed as *particles* of the state variables or parameters. These weighted samples are used to approximate a complicated or a non-Gaussian Bayes' *a posteriori* PDF. There have also been many efficient modifications and improvements on these methods briefly described in Chapter: 2.

In this chapter, a brief review of the generic PF also known as *bootstrap* PF or (SIS-R) filter, and parametric PFs based on Gaussian and GMM approximation of Bayes' *a posteriori* PDF will be carried out. Next, new efficient SMC methods are developed that utilize the GCS and its mixture model to augment and improve the standard PF. The filtering methods include GCS Particle Filter (GCSPF), GCS Mixture Particle Filter (GCSMPF), and Hybrid GCS Culver Particle Filter (HGCPF). The first algorithm, GCSPF is an extension of Gaussian PF (GPF) by reference [42] and the last algorithm HGCPF is nonlinear MC adaptation and modification of Culver Filter (CF) [1] .

There are situations where the evolution of a dynamical system cannot be measured at each time instant, for example, in space object (i.e., satellites or space debris) radar tracking requirements, the physical appearance of an object over the horizon is needed to record radar measurements (details later in this chapter). The appearance is usually 5-10 minutes for a LEO object depending upon a particular type of orbit. This forms about $1/10^{\text{th}}$ of the time taken by the object to orbit around the Earth. Moreover, there could be practical limitations associated with measurement devices, which restrict availability of measurements at each time instant during the appearance as well. Therefore, an ability to accurately predict the state evolution for such a dynamical system along with state uncertainty i.e., state probability distribution (state predictive PDF) is very critical. The filtering

algorithms based on Gaussian or GMM approximation of state predictive PDF such as EKF [6], GPF [42] or GSPF [43] may not be sufficient for such requirements. In this chapter, two nonlinear dynamical systems modelled in continuous time form i.e., a simple pendulum and satellite orbital dynamics have been used for the implementation of filtering algorithms discussed in the chapter, under less or sparse measurements availability. The simple pendulum's analogy to various two dimensional nonlinear physical phenomena has lead researchers in filtering community to experiment their filtering algorithm [21]. For example consider the cross track relative motion of a satellite described by HCW Equations (see third of Equation: 3.49), which can be considered as a simple harmonic oscillator similarly to the simple pendulum. The particle filtering algorithms based on GCS and its mixture models have shown improved performances over other methods.

5.2 Fundamentals of Particle Filters

Particle filtering is based on MC simulations to obtain approximation of PDFs given in Equations: 5.3 to 5.7. The main objective is to sequentially sample and resample particles from a particular choice of PDF known as *proposal* PDF, considered by the filter as approximation of Bayes' *a posteriori* PDF. The choice of *proposal* PDF is a major issue for the different variants of PF [22],[13],[42],[45]. Optimal Bayesian estimation (Equations: 5.3 to 5.7) is directly implemented, wherein entire Bayes' *a posteriori* PDF is approximated sequentially.

5.2.1 Monte Carlo Integration

PFs employ MC integration scheme to compute integrals. For example, an ensemble of weighted particles (samples), acquired from Bayes' *a posteriori* PDF can be used to formulate integrals into discrete sums. Therefore, one may approximate Bayes' *a posteriori* PDF as [22],[47]:

$$p(\mathbf{x}_k | \mathbf{y}_{1:k}) \approx \hat{p}(\mathbf{x}_k | \mathbf{y}_{1:k}) = \frac{1}{N} \sum_{i=1}^N \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}) \quad (5.8)$$

where, randomly distributed samples $\{\mathbf{x}_k^{(i)}; i = 1, 2, \dots, N\}$, are drawn from $p(\mathbf{x}_k | \mathbf{y}_{1:k})$, N is the number of samples, and $\delta(\cdot)$ denotes the Dirac delta function [104].

Therefore, any expectations of form expressed in Equation: 5.3 can be approximated by the following estimates:

$$E[g(\mathbf{x}_k)] \approx \hat{E}[g(\mathbf{x}_k)] = \frac{1}{N} \sum_{i=1}^N g(\mathbf{x}_k^{(i)}) \quad (5.9)$$

Assume that $\{\mathbf{x}_k^{(i)}; i = 1, 2, \dots, N\}$ are independent variables and each random variable has the same marginal PDF [99]. As a consequence of being from the same marginal PDF the variables are said to be identically distributed. Therefore, if the particles $\mathbf{x}_k^{(i)}$ are independent and identically distributed (i.i.d) then its mean can be computed as follows [99]:

$$\hat{\mathbf{x}}_k = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_k^{(i)} \quad (5.10)$$

The covariance $\hat{\mathbf{P}}_k$ can be approximated by:

$$\hat{\mathbf{P}}_k = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_k^{(i)} - \hat{\mathbf{x}}_k) (\mathbf{x}_k^{(i)} - \hat{\mathbf{x}}_k)^T \quad (5.11)$$

According to law of large numbers as N approaches infinity the estimates and true expectations converge almost surely [22].

$$\hat{E}[g(\mathbf{x}_k)] \xrightarrow{N \rightarrow \infty} E[g(\mathbf{x}_k)] \quad (5.12)$$

where, $\hat{E}[g(\mathbf{x}_k)]$ is MC expectation (Equation: 5.9).

5.2.2 Bayesian Importance Sampling

One can approximate the Bayes' *a posteriori* PDF with a discrete function as shown in Equation: 5.8. However, samples cannot be drawn from this PDF as it is not known. One may overcome this problem by sampling from a known, easy to sample, *proposal* PDF $p(\mathbf{x}_k|\mathbf{y}_{1:k})$. This procedure is known as *importance* sampling [22]. The selection of this distribution is an important design issue for different variants and/or improvements of particle based inference algorithms like Extended Kalman Particle Filter (EKPF) [105], Sigma Point Particle Filter (SPPF), and Gaussian Mixture Sigma Point Particle Filter (GMSPPF) [22]. Expectations for functions of states (Equation: 5.9) are computed from particles drawn from *proposal* PDF. For example $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ could be a PDF with a complex function or no analytical expression and $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ could be an analytical Gaussian PDF. Therefore, one can write $p(\mathbf{x}_k|\mathbf{y}_{1:k}) \propto q(\mathbf{x}_k|\mathbf{y}_{1:k})$ where, the symbol \propto means that $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ is proportional to $q(\mathbf{x}_k|\mathbf{y}_{1:k})$ at every \mathbf{x}_k . As $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ is normalized density function, then $q(\mathbf{x}_k|\mathbf{y}_{1:k})$ must be scaled un-normalized equivalent of $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ with a unique scaling weight at each \mathbf{x}_k [103]. Thus we may write scaling factor or weight as [22],[8]:

$$w_k(\mathbf{x}_k) = \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{p(\mathbf{x}_k|\mathbf{y}_{1:k})} \quad (5.13)$$

Now the MC Expectation of $g(\mathbf{x}_k)$ (by making use Equation: 5.13) can be derived as (for proof see ref [22]):

$$\begin{aligned} E[g(\mathbf{x}_k)] &= \int_{-\infty}^{+\infty} g(\mathbf{x}_k) \frac{p(\mathbf{x}_k|\mathbf{y}_{1:k})}{p(\mathbf{x}_k|\mathbf{y}_{1:k})} p(\mathbf{x}_k|\mathbf{y}_{1:k}) d\mathbf{x}_k \\ &= \frac{E_p[g(\mathbf{x}_k)w_k(\mathbf{x}_k)]}{E_p[w_k(\mathbf{x}_k)]} \\ E[g(\mathbf{x}_k)] &\approx \hat{E}_p[g(\mathbf{x}_k)] = \frac{\frac{1}{N} \sum_{i=1}^N g(\mathbf{x}_k^{(i)}) w_k(\mathbf{x}_k^{(i)})}{\frac{1}{N} \sum_{i=1}^N w_k(\mathbf{x}_k^{(i)})} \end{aligned} \quad (5.14)$$

If one samples N particles from $p(\mathbf{x}_k|\mathbf{y}_{1:k})$ then the expectation of interest can now easily be re-expressed using particle representation of *proposal* PDF as:

$$E[g(\mathbf{x}_k)] \approx \hat{E}[g(\mathbf{x}_k)] = \sum_{i=1}^N \tilde{w}_k^{(i)} g(\mathbf{x}_k^{(i)}) \quad (5.15)$$

where, the normalized weights $\tilde{w}_k^{(i)}$ are given by:

$$\tilde{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{j=1}^N w_k^{(j)}} \quad (5.16)$$

5.2.3 Sequential Importance Sampling

In order to obtain sequential state estimates one has to construct a sequential form of the *proposal* PDF for sampling and use the Equation: 5.15. Let $\mathbf{x}_k = (\mathbf{x}_0^T, \mathbf{x}_1^T \dots \mathbf{x}_k^T)^T$ and $\mathbf{y}_k = (\mathbf{y}_1^T, \mathbf{y}_2^T \dots \mathbf{y}_k^T)^T$ be the stacked vector of states and observations up to time step k . Under the assumption of the state being a Markov process [5] one may write [8] :

$$p(\mathbf{x}_{0:k}) \triangleq p(\mathbf{x}_k) = p(\mathbf{x}_0) \prod_{i=1}^k p(\mathbf{x}_i|\mathbf{x}_{i-1}) \quad (5.17)$$

The measurements \mathbf{y}_k are considered as conditionally independent, given the states \mathbf{x}_k :

$$p(\mathbf{y}_{1:k}|\mathbf{x}_{0:k}) \triangleq p(\mathbf{y}_{1:k}|\mathbf{x}_k) = \prod_{i=1}^k p(\mathbf{y}_i|\mathbf{x}_i) \quad (5.18)$$

One can now conveniently express a sequential form of *proposal* PDF on the basis of above Markov property [5] of DSSM (Equations: 5.17 and 5.18) as [47]:

$$\begin{aligned} \mathcal{p}(\mathbf{x}_{0:k}|\mathbf{y}_{1:k}) &= \mathcal{p}(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k}) \times \mathcal{p}(\mathbf{x}_0) \prod_{i=1}^{k-1} \mathcal{p}(\mathbf{x}_i|\mathbf{y}_i) \\ \mathcal{p}(\mathbf{x}_{0:k}|\mathbf{y}_{1:k}) &= \mathcal{p}(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k}) \mathcal{p}(\mathbf{x}_{0:k-1}|\mathbf{y}_{1:k-1}) \end{aligned} \quad (5.19)$$

Similarly, the weight equation (Equation: 5.13) can be re-expressed in terms of full states and measurements up to time “ k ” as:

$$w_k(\mathbf{x}_{0:k}) = \frac{p(\mathbf{y}_{1:k}|\mathbf{x}_{0:k})p(\mathbf{x}_{0:k}|\mathbf{y}_{1:k-1})}{\mathcal{p}(\mathbf{x}_{0:k}|\mathbf{y}_{1:k})} \quad (5.20)$$

By substituting Equations: 5.17 to 5.19 in Equation: 5.20, recursive estimates for weights can be expressed as (for proof see [22]):

$$w_k = w_{k-1} \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})}{\mathcal{p}(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k})} \quad (5.21)$$

The most popular choice for *proposal* PDF expressed in the denominator of Equation: 5.21, is the state transition PDF $p(\mathbf{x}_k|\mathbf{x}_{k-1})$, primarily due to ease of implementation [41],[22]:

$$\mathcal{p}(\mathbf{x}_k|\mathbf{x}_{0:k-1}, \mathbf{y}_{1:k}) = p(\mathbf{x}_k|\mathbf{x}_{k-1}) \quad (5.22)$$

By substituting Equation: 5.22 into Equation 5.21, the recursive weight expression becomes:

$$w_k = w_{k-1} p(\mathbf{y}_k|\mathbf{x}_k) \quad (5.23)$$

Equation: 5.23 can easily be implemented by obtaining sample from the state transition PDF

$p(\mathbf{x}_k|\mathbf{x}_{k-1})$, evaluation of the measurement likelihood $p(\mathbf{y}_k|\mathbf{x}_k)$, which will then be multiplied by sample weights from the previous time step ($k-1$). The recursive flow of weights starts by generating an initial set of particles (samples) of equal weight:

$$w_0^{(i)} = \frac{1}{N}, i = 1 \dots N \quad (5.24)$$

This procedure is known as Sequential Importance Sampling (SIS) [22].

5.2.4 Degeneration of Particles and its Minimization

Sequential estimation would require a repeated use and sampling from the state transition PDF and implementation of Equation: 5.23. However, the disadvantage of this simple approach is dispersion of particles from the expected value of $\hat{\mathbf{x}}_k$ due to unbounded increase of variance of \mathbf{x}_k as $k \rightarrow \infty$ [103]. Thus, the sample $\mathbf{x}_k^{(i)}$ that disperses from the expected value of $\hat{\mathbf{x}}_k$, its weight $w_k^{(i)}$ approaches zero. This problem has been termed as *degeneracy* of particle filters. To measure the degeneracy of the particle filter, the effective sample size, N_E is computed. It is a way to measure how well particles are concentrated in the regions of interest and is expressed as [64]:

$$N_E = \frac{1}{\sum_{i=1}^N (w_k^{(i)})^2} \quad (5.25)$$

Degeneration of particles is highly undesirable. To reduce its effect, one may employ a brute force method of increasing the number of particles for filtering at the cost of prohibitively high computation. Another approach to minimize the effect of this problem is resampling of particles [8],[47],[41]. Resampling is essentially elimination of samples with low importance weights and multiplication of samples with high importance weights [22]. In this step one generate children samples N_i associated to each particle $\mathbf{x}_k^{(i)}$ such that, $\sum_{i=1}^N N_i = N$. Different types of resampling techniques are proposed such as SIR [45] and Residual Resampling (RR) [64],[22],[105]. In SIR a Dirac random measure $\{\mathbf{x}_k^{(i)}, \tilde{w}_k^{(i)}\}$ is mapped into an equally weighted random measure $\{\mathbf{x}_k^{(j)}, 1/N\}$. This is accomplished by sampling from a discrete set $\{\mathbf{x}_k^{(i)}; i = 1, 2, \dots, N\}$ with probabilities $\{\tilde{w}_k^{(i)}; i = 1, 2, \dots, N\}$. Firstly, the Cumulative Distribution Function (CDF) using the weights $\tilde{w}_k^{(i)}$ is constructed. Then one obtains the sampling index (i) from the uniform distribution ($\mathcal{U}[0,1]$) and projects it onto the distribution range and then onto the distribution domain (see Figure: 5-3). The intersection with the domain constitutes the new resampled index (j). That means that a particle $\mathbf{x}_k^{(j)}$ is selected as a new sample. Therefore, the particles with larger weights will end up having more children [105].

RR is performed in two steps. In the first step the number of particles / children are deterministically computed using the floor function ($\lfloor(\cdot)\rfloor$):

$$\tilde{N}_i = \left\lfloor N \tilde{w}_k^{(i)} \right\rfloor \quad (5.26)$$

where, each $\mathbf{x}_k^{(i)}$ particle is replicated \tilde{N}_i times. In the second step, SIR is used to select the remaining \bar{N}_k particles:

$$\bar{N}_k = N - \sum_{i=1}^N \tilde{N}_i \quad (5.27)$$

with new weights expressed for each particle as:

$$\tilde{w}_k^{(i)} = \bar{N}_k^{-1} \left(\tilde{w}_k^{(i)} N - \tilde{N}_i \right) \quad (5.28)$$

The children samples for each individual particle (as obtained from SIR) would form the second set \tilde{N}_i , such that $\bar{N}_k = \sum_{i=1}^N \tilde{N}_i$. Finally the results are added to get the total number of children for each sample $N_i = \tilde{N}_i + \tilde{N}_i$. In general, an adaptive resampling strategy is adopted in the PF wherein the resampling step is only performed if effective size of particles N_E (Equation: 5.25) becomes less than some threshold size N_T .

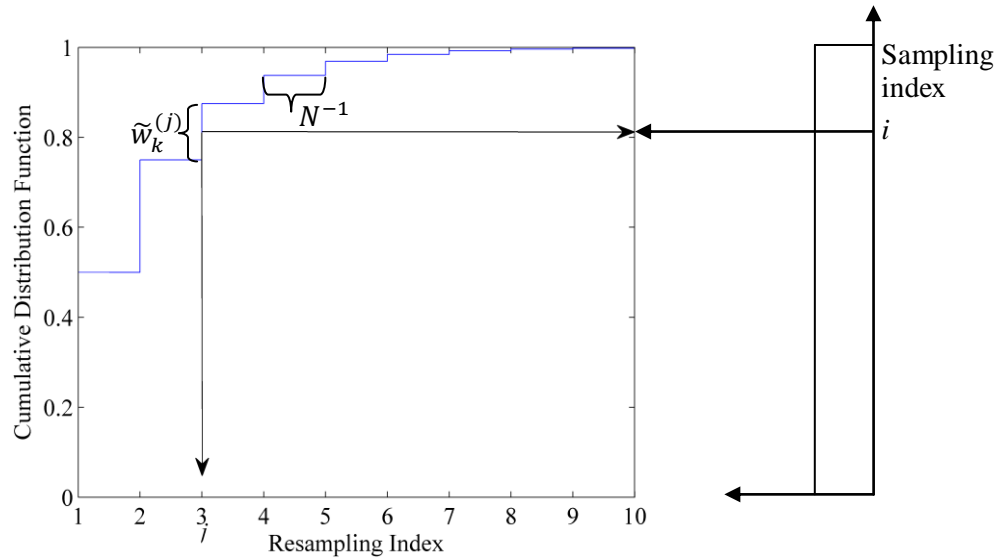


Figure 5-3: In SIR a random measure $\{\mathbf{x}_k^{(i)}, \tilde{w}_k^{(i)}\}$ is mapped into equally weighted random measure $\{\mathbf{x}_k^{(j)}, N^{-1}\}$. The index i is drawn from a uniform distribution shown on right hand side (not to scale).

5.2.5 Generic Bootstrap Particle Filter Algorithm

As expressed in Equation: 5.22, the state transition PDF is the most popular choice of *proposal* PDF due to the ease of implementation. For this choice of *proposal* PDF the generic PF is also known as SIS-R or *bootstrap* PF [41]. The term *bootstrap* is associated due to non parametric form of the PDF approximation by samples [67]. Successful implementations of the SIS-R algorithm assume: (1) availability of a suitable *proposal* PDF for sampling and resampling and, (2) Dirac point mass approximation of the Bayes' *a posteriori* PDF. Therefore, if these conditions are not met the PF may produce undesirable estimates. One may increase the number of particles for filtering but it requires heavy computations. Moreover, in order to capture the true structure of the Bayes' *a posteriori* PDF, which might be multi-modal, sample variety is highly desirable. A resampling stage may introduce depletion of samples, therefore; it is unable to form approximation of true Bayes' *a posteriori* PDF with sufficient accuracy due to the multiple duplication of the same sample with a higher weight. Thus samples might eventually collapse, to a single sample with most dominant weight. This situation would severely degenerate PF output.

5.2.6 Parametric Bootstrap Particle Filtering Algorithms

In parametric *bootstrap* based PFs, assumptions on the form of the Bayes' *a posteriori*, state predictive and transition PDFs is considered. In this section a brief review on the work by reference [42],[43], would be considered. This is based on the assumption of Gaussian or GMM form for aforementioned PDFs. The PF which is based on single Gaussian PDF is known as Gaussian Particle Filter (GPF) [42] and the one based on GMM is Gaussian Sum Particle Filter (GSPF) [43].

5.2.6.1 Gaussian Particle Filter

The GPF approximates the state predictive and Bayes' *a posteriori* PDF as Gaussian. However, contrary to the EKF, which also assumes that these PDFs are Gaussian, and employ linearization of the functions in the process and observation equations (Equation: 5.1), the GPF generates the Gaussian approximations by using particles that are propagated through process and observation equations without approximation. At " k^{th} " instant of time, the samples obtained from initial state PDF $p(\mathbf{x}_{k-1})$ are propagated forward in time (referred as *time update*) through the nonlinear function as expressed in Equation: 5.1. This would provide particle approximation of the state transition PDF $p(\mathbf{x}_k|\mathbf{x}_{k-1})$. An MC integration is performed to obtain the mean and covariance of the state predictive PDF using the following equation [42]:

$$p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \int_{-\infty}^{+\infty} p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{1:k-1})p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1})d\mathbf{x}_{k-1} \quad (5.29)$$

$$p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) = \frac{1}{N} \sum_{i=1}^N p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_{1:k-1})$$

The sample mean and covariance for state predictive PDF $p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$ are expressed as follows:

$$\begin{aligned} \boldsymbol{\mu}_k &= \frac{1}{N} \sum_{i=1}^N \mathbf{x}_k^{(i)} \\ \mathbf{P}_k &= \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \end{aligned} \quad (5.30)$$

where, the particles $\mathbf{x}_k^{(i)}$ are obtained from state transition PDF $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{1:k-1})$

In the *measurement update* (Bayesian update using Equation: 5.4) resampling of particles from state predictive PDF is performed. These particles are then used to compute weights by evaluating measurement likelihood $p(\mathbf{y}_k | \mathbf{x}_k)$ as in the PF. The weights $w_k^{(i)}$ are computed in a non-iterative manner using the following likelihood:

$$w_k^{(i)} = p(\mathbf{y}_k | \mathbf{x}_k^{(i)}) \quad (5.31)$$

This is followed by the normalization step given in Equation: 5.16. The inference of the mean and covariance is then drawn using these normalized weights. Resampling techniques i.e., residual resampling described earlier are not required for the GPF. Unlike SIS-R PF, GPF computes weights in non-sequential manner using Equation: 5.31. Better choice of proposal density are possible in GPF i.e., EKF generated Bayes' *a posteriori* PDF. However, this needs a separate EKF running in parallel which makes it susceptible to linearization errors.

5.2.6.2 Gaussian Sum Particle Filter

Any probability density $p(\mathbf{x}_k)$ can be approximated as closely as desired by a GMM of the following form [22],[60]:

$$p(\mathbf{x}_k) \approx p_{gmm}(\mathbf{x}_k) = \sum_{g=1}^G \alpha_k^{(g)} \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}) \quad (5.32)$$

where, G is the number of mixing components, $\alpha_k^{(g)}$ are the mixing weights of g^{th} component and $\mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)})$ denotes the Gaussian (normal) PDF function with $\boldsymbol{\mu}_k^{(g)}$ is the mean vector of g^{th}

component, and $\mathbf{P}_k^{(g)}$ is the positive definite covariance matrix of g^{th} component.

Therefore, one may write the GMM Bayes' *a posteriori* and process noise PDFs at $(k - 1)$ time instant as:

$$p_{gmm}(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1}) = \sum_{g=1}^G \alpha_{k-1}^{(g)} \mathcal{N}(\mathbf{x}_{k-1}, \boldsymbol{\mu}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{(g)}) \quad (5.33)$$

$$p_{gmm}(\mathbf{w}_{k-1}) = \sum_{i=1}^I \beta_{k-1}^{(i)} \mathcal{N}(\mathbf{w}_{k-1}, \boldsymbol{\mu}_{\mathbf{w},k-1}^{(i)}, \mathbf{Q}_{k-1}^{(i)}) \quad (5.34)$$

where, \mathbf{w}_{k-1} is the process noise variable with $\boldsymbol{\mu}_{\mathbf{w},k-1}$ being the mean vector and \mathbf{Q}_{k-1} is the covariance of the process noise, respectively. Now in order to initialize the filter for the *time update* (prediction of dynamics without using measurements), consider the availability of GMM approximated Bayes' *a posteriori* PDF at time $(k - 1)$ (Equation: 5.33). The goal is to obtain the state the state predictive PDF $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{k-1})$ also as GMM. As already defined for a single Gaussian PDF (Equation: 5.29), the state transition PDF $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is now defined in terms of a probabilistic model governing the system's state evolution and process noise statistics i.e., GMM. The state transition PDF can be expressed as [43]:

$$p_{gmm}(\mathbf{x}_k|\mathbf{x}_{k-1}) = \sum_{i=1}^I \beta_{k-1}^{(i)} \mathcal{N}(\mathbf{x}_k; \mathbf{f}(\mathbf{x}_{k-1}), \mathbf{Q}_{k-1}^{(i)}) \quad (5.35)$$

where, $\mathbf{f}(\cdot)$ is the nonlinear process model expressed in Equation: 5.1. For the sake of simplicity consider $\boldsymbol{\mu}_{\mathbf{w},k-1} = 0$. After substituting Equation: 5.33 and 5.35 in Equation: 5.5 one has:

$$\begin{aligned} p_{gmm}(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{k-1}) \\ = \sum_{g=1}^G \sum_{i=1}^I \beta_{k-1}^{(i)} \alpha_{k-1}^{(g)} \int_{-\infty}^{+\infty} \mathcal{N}(\mathbf{x}_k; \mathbf{f}(\mathbf{x}_{k-1}), \mathbf{Q}_{k-1}^{(i)}) \mathcal{N}(\mathbf{x}_{k-1}, \boldsymbol{\mu}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{(g)}) d\mathbf{x}_{k-1} \end{aligned} \quad (5.36)$$

The expression inside the integral in Equation: 5.36, is quite extensive and may not be solvable due to nonlinearity of the process equation (Equation: 5.1) [43]. However, the solution of this integral can be approximated by the Gaussian PDF [60]. Therefore, using the similar procedure as adopted for a single Gaussian PDF in GPF (Equation: 5.29 and 5.30), is now being used here for each individual GMM component separately. This provides the GMM approximation of the state predictive PDF. In

the *measurement update* (Bayesian update using Equation: 5.4) step, resampling from the state predictive PDF is performed, in order to compute weights for each sample of individual GMM component, using the measurement likelihood $p(\mathbf{y}_k|\mathbf{x}_k)$. Due to the repeated use of Equation: 5.36 the size of the number of components of GMM Bayes' *a posteriori* PDF would grow exponentially. However, this can be resolved by using a resampling step i.e., RR in *measurement update*. The small weights are discarded whereas children samples are produced for GMM components having high weights. The subsequent sections develop more efficient particle filtering algorithms based on GCS and its mixture model. Hence, SIS-R PF, GPF and GSPF will be used for comparison purposes.

5.3 Gram Charlier Series

GCS is an orthogonal series expansion of a PDF in terms of its higher order moments. It can be utilized to approximate arbitrary PDFs, especially heavy tails and any higher order PDF structures like skew and kurtosis [28],[26]. It is a very rich classical form similar to Taylor series and is based on the Gaussian PDF, developed in early 19th Century by [29],[32]. The series employs a set of Hermite polynomials which are orthogonal with respect to a Gaussian weighting function i.e., $e^{-1/2x^2}$ over the domain $(-\infty, \infty)$ [31].

5.3.1 Univariate GCS

The univariate GCS expansion $p_{gcs}(x_k)$ of an arbitrary PDF $p(x_k)$ around its best Gaussian estimate $p_g(x_k)$ with mean (μ), and standard deviation (σ), is given by [28],[34]:

$$p(x_k) \approx p_{gcs}(x_k) \quad (5.37)$$

$$p_{gcs}(x_k) = \mathcal{N}(x_k, \mu, \sigma) \left[1 + \frac{1}{3!} \kappa_3 h_3(x_k, \mu, \sigma) + \frac{1}{4!} \kappa_4 h_4(x_k, \mu, \sigma) \right. \\ \left. + \frac{1}{5!} \kappa_5 h_5(x_k, \mu, \sigma) + \frac{10}{6!} \kappa_6 h_6(x_k, \mu, \sigma) + \dots \right]$$

where, κ_i is the i^{th} standardized cumulant (defined as $\kappa_i = \frac{\kappa_i}{\sigma^i}$) and h_i is the univariate Hermite polynomial of order i . The standard Hermite polynomials of order n can be obtained by putting $\mu = 0$ and $\sigma^2 = 1$ using Rodrigues formula expressed as [27]:

$$h_n(z) = (-1)^n e^{\frac{z^2}{2}} \frac{d^n}{dz^n} e^{-\frac{z^2}{2}} \quad (5.38)$$

where, $z = \frac{(x-\mu)}{\sigma}$.

The Hermite polynomials obey the following recursive relationship [26]:

$$\frac{d}{dz} h_n(z) = n h_{n-1}(z) \quad (5.39)$$

5.3.2 Multivariate GCS

In a manner similar to the orthogonally expanded univariate PDFs, multivariate expansions can be described. If all the moments of a d -dimensional random vector \mathbf{x}_k are finite, then any probability density $p(\mathbf{x}_k)$ can be expressed as Gaussian density $\mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k)$ multiplied by an infinite series of multidimensional Hermite polynomials as [28]:

$$\begin{aligned} p_{gcs}(\mathbf{x}_k) = \mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) & \left[1 + \sum_{i,j,l} \frac{\kappa_{i,j,l}}{3!} h_{ijl}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) \right. \\ & + \sum_{i,j,l,m} \frac{\kappa_{i,j,l,m}}{4!} h_{ijlm}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) \\ & + \sum_{i,j,l,m,n} \frac{\kappa_{i,j,l,m,n}}{5!} h_{ijlmn}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) \\ & \left. + \sum_{i,j,l,m,n,o} \frac{(\kappa_{i,j,l,m,n,o} + \kappa_{i,j,l} \kappa_{m,n,o} [10])}{6!} h_{ijlmno}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) + \dots \right] \end{aligned} \quad (5.40)$$

where, the subscripts i, j, l denotes the dimension, k is the time subscript. The functions $h_{i\dots o}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k)$ and similar forms are multidimensional Hermite polynomials with corresponding input dimensions $i, j, l \in \{1, \dots, d\}$, and $\kappa_{i,j,l}$ is the corresponding third multivariate cumulant over input dimensions i, j, l , where sum over all input dimensions i, j, l is considered. Similarly, $\kappa_{i,j,l,m}$ is the fourth multivariate cumulant and $\kappa_{i,j,l,m,n}$ is the fifth multivariate cumulant and time subscript (k) are omitted (considering their time dependence implicitly) for multivariate cumulants to simplify their notation. Hermite polynomials can be obtained by differentiating $\mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k)$ again using the Rodrigues formula [33]:

$$\begin{aligned} h_{i\dots o}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) & \\ & = (-1)^{i+\dots+o} \frac{1}{\mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k)} \frac{d^{i+\dots+o}}{dx_1^i \dots dx_d^o} \mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) \end{aligned} \quad (5.41)$$

Some useful functional forms of Hermite polynomials are expressed as [28]:

$$\begin{aligned}
h_i(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) &= P_{ij}^{-1}(x_j - \mu_j) \\
h_{ij}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) &= h_i h_j - P_{ij}^{-1} \\
h_{ijl}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) &= h_i h_j h_l - h_i P_{jl}^{-1}[3] \\
h_{ijlm}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) &= h_i h_j h_l h_m - h_i h_j P_{lm}^{-1}[6] + P_{ij}^{-1} P_{lm}^{-1}[3] \\
h_{ijklmn}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k) &= h_i h_j h_l h_m h_n - h_i h_j h_l P_{mn}^{-1}[10] + h_i P_{jl}^{-1} P_{mn}^{-1}[15]
\end{aligned} \tag{5.42}$$

where, P_{ij}^{-1} and similar forms indicate the ij^{th} component of the inverse of covariance matrix \mathbf{P}_k , x_j and μ_j indicate the j^{th} variable and its mean, respectively. The subscripts implicitly imply summation over indices. The connection between cumulants and multivariate central moments is defined as [28]:

$$\begin{aligned}
\kappa_{i,j} &= P_{ij} \\
\kappa_{i,j,l} &= P_{ijl}^{(3)} \\
\kappa_{i,j,l,m} &= P_{ijlm}^{(4)} - P_{ij} P_{lm}[3] \\
\kappa_{i,j,l,m,n} &= P_{ijlmn}^{(5)} - P_{ij} \kappa_{l,m,n}[10] \\
\kappa_{i,j,l,m,n,o} &= P_{ijlmno}^{(6)} - P_{ij} \kappa_{l,m,n,o}[15] - \kappa_{i,j,l} \kappa_{m,n,o}[10] \\
&\quad - P_{ij} P_{lm} P_{no}[15]
\end{aligned} \tag{5.43}$$

where, $P_{ijl}^{(3)}$ and similar forms indicate the ijl^{th} component of the third order (coskewness) tensor etc. The bracket notations used in Equations: 5.40, 5.42 and 5.43 are sums over partition of combinations of indices. For example:

$$P_{ij} P_{lm}[3] = P_{ij} P_{lm} + P_{il} P_{jm} + P_{im} P_{jl}$$

5.4 Gram Charlier Series Mixture Model

A detailed viewpoint on single GCS expansions has already been described earlier in Chapter: 1. The GCS expansions of lower order (order ≤ 4) do not estimate well near the centroid of the PDF. Moreover, the resulting PDF could be negative and not unimodal [106]. To improve the density estimation accuracy one can increase the order of these expansions, but unfortunately it renders the estimate more sensitive to outliers. Therefore, rather than increasing the order of the GCS expansion

by using single Gaussian PDF (Equation: 5.37 and 5.40), it was suggested by [34] to use mixtures of GCS expanded Gaussian PDFs of moderate the order. Therefore, one now describes GCSMM and later in this chapter, this form of orthogonal expansion is considered for improving particle filters for nonlinear dynamical systems.

5.4.1 Univariate Gram Charlier Series Mixture Model

The univariate GCSMM approximation $p_{gcsm}(x_k)$ until order four, for an arbitrary non-Gaussian PDF $p(x_k)$ is given by [28],[34]:

$$p(x_k) \approx p_{gcsm}(x_k) \quad (5.44)$$

$$p_{gcsm}(x_k) = \sum_{g=1}^G \alpha_k^{(g)} \mathcal{N}(x_k, \mu_k^{(g)}, \sigma_k^{(g)}) \left[1 + \frac{1}{3!} \kappa_3^{(g)} h_3(x_k, \mu_k^{(g)}, \sigma_k^{(g)}) + \frac{1}{4!} \kappa_4^{(g)} h_4(x_k, \mu_k^{(g)}, \sigma_k^{(g)}) \right]$$

where, G is the number of GCSMM components. The parameters of above mixture PDF can be conveniently estimated using statistical Expectation Maximization (EM) Algorithm [34],[67],[107],[108] (more details of EM are described in Section: 5.4.2) :

$$\alpha^{(g)} = \frac{1}{N} \sum_j \tau_j^{(g)}, \mu_1^{(g)} = \frac{1}{N} \sum_j \frac{\tau_j^{(g)} x_j}{\alpha^{(g)}}, \mu_2^{(g)} = \frac{1}{N} \sum_j \frac{\tau_j^{(g)} (x_j - \mu_1^{(g)})^2}{\alpha^{(g)}}, \quad (5.45)$$

$$\mu_3^{(g)} = \frac{1}{N} \sum_j \frac{\tau_j^{(g)} (x_j - \mu_1^{(g)})^3}{\alpha^{(g)}}, \mu_4^{(g)} = \frac{1}{N} \sum_j \frac{\tau_j^{(g)} (x_j - \mu_1^{(g)})^4}{\alpha^{(g)}},$$

where,

$$\tau_j^{(g)} = \frac{\alpha^{(g)} \mathcal{N}(x_k, \mu_k^{(g)}, \sigma_k^{(g)}) \left[1 + \frac{1}{3!} \kappa_3^{(g)} h_3(x_k, \mu_k^{(g)}, \sigma_k^{(g)}) + \frac{1}{4!} \kappa_4^{(g)} h_4(x_k, \mu_k^{(g)}, \sigma_k^{(g)}) \right]}{\sum_{g=1}^G \alpha_k^{(g)} \mathcal{N}(x_k, \mu_k^{(g)}, \sigma_k^{(g)}) \left[1 + \frac{1}{3!} \kappa_3^{(g)} h_3(x_k, \mu_k^{(g)}, \sigma_k^{(g)}) + \frac{1}{4!} \kappa_4^{(g)} h_4(x_k, \mu_k^{(g)}, \sigma_k^{(g)}) \right]}$$

j is the subscript for j^{th} data point, N is the number of data points, $\tau_j^{(g)}$ is the *posterior* probability and $\mu_1, \mu_2, \mu_3, \mu_4$ are the mean, the second, third and fourth order univariate moments respectively (for proof of higher order EM equations for univariate moments μ_3 and μ_4 in Equation: 5.45 see [34]).

The standardized third and fourth cumulants are [26]:

$$\kappa_3^{(g)} = \frac{\mu_3^{(g)}}{\sqrt{\mu_2^{(g)} \mu_2^{(g)} \mu_2^{(g)}}}, \quad (5.46)$$

$$\kappa_4^{(g)} = \frac{\mu_4^{(g)} - 3\mu_2^{(g)} \mu_2^{(g)}}{\sqrt{\mu_2^{(g)} \mu_2^{(g)} \mu_2^{(g)} \mu_2^{(g)}}}.$$

The quantities computed in Equation: 5.45 and 5.46 are called parameters of a univariate GCSMM. As an example, consider the GCSMM in Equation: 5.44 to approximate two non-Gaussian PDF's i.e., the exponential and uniform and compare them with a single GCS and GMM approximations. The comparison is illustrated in Figures: 5-4 and 5-5. The figures clearly indicate inability of single GCS to capture centroid of the non-Gaussian PDFs. Moreover, if a single GCS is truncated at lower order of Hermite polynomial, then it might produce negative probability regions. Negative probability regions are visible in both these figures. Table: 5-1 presents the Root Mean Square Error (RMSE) for these approximations. The RMSE clearly suggests improvement in approximating non-Gaussian PDFs using mixture models and one could consider the GCSMM as a natural extension to the GMM.

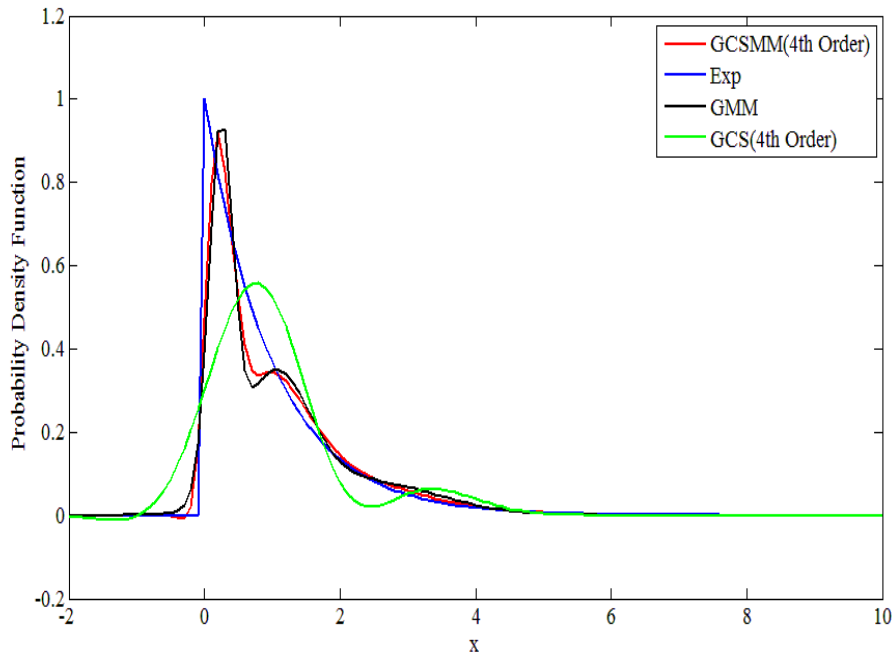


Figure 5-4: The comparison of true exponential PDF with GCSMM ($3 \times$ components), GMM ($3 \times$ components) and single GCS approximation.

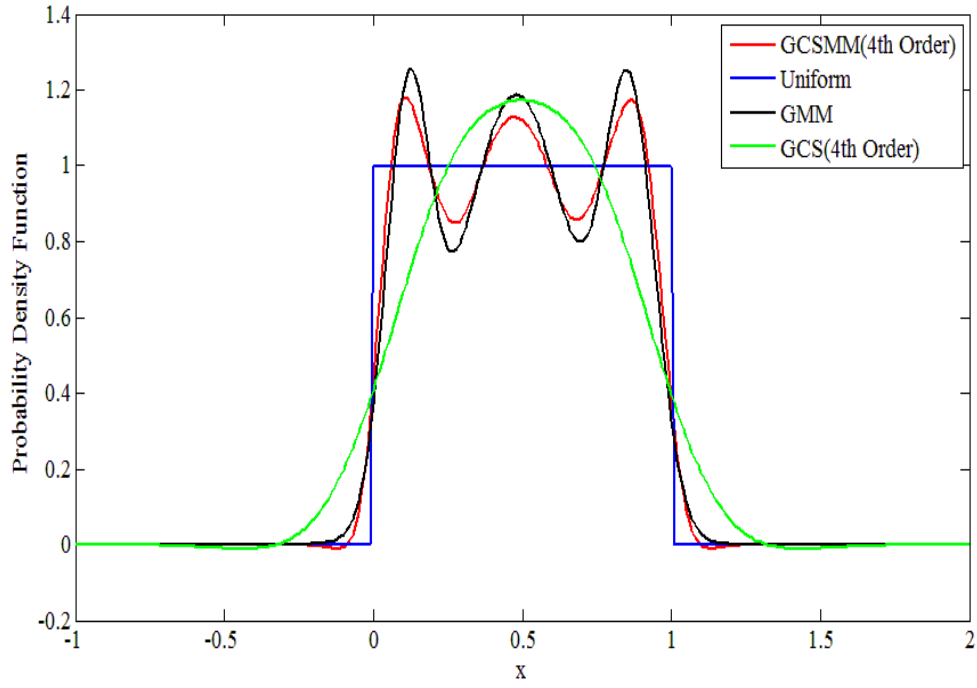


Figure 5-5: The comparison of true uniform PDF with GCSMM ($3 \times$ components), GMM ($3 \times$ components) and single GCS approximations.

PDF	Exponential	Uniform
GCS	0.109070	0.171282
GMM	0.071867	0.133243
GCSMM	0.057867	0.106224

Table 5-1: RMSE PDF approximations comparison results

5.4.2 Multivariate GCSMM

On similar lines to univariate GCSMM, one can approximate a d -dimensional arbitrary non-Gaussian PDF $p(\mathbf{x}_k)$ using a mixture of multivariate GCS, as expressed in Equation: 5.40. The GCSMM expansion of $p(\mathbf{x}_k)$ up to order five is given by:

$$\begin{aligned}
p_{gcsm}(\mathbf{x}_k) = & \sum_{g=1}^G \alpha_k^{(g)} \mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}) \left[1 + \sum_{i,j,l} \frac{\kappa_{i,j,l}^{(g)}}{3!} h_{ijl}(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}) \right. \\
& + \sum_{i,j,l,m} \frac{\kappa_{i,j,l,m}^{(g)}}{4!} h_{ijlm}(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}) \\
& \left. + \sum_{i,j,l,m,n} \frac{\kappa_{i,j,l,m,n}^{(g)}}{5!} h_{ijlmn}(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}) \right], \tag{5.47}
\end{aligned}$$

where, G is the number of mixands.

The parameters of the above PDF including higher order moments and cumulants can be readily estimated by adapting the EM Algorithm [107],[108],[22] for a GMM. The details of these adjustments are now being described.

Essentially, the parameters of the multivariate GCSMM are estimated by adapting the concept of parameter estimates of univariate GCSMM (Equation: 5.45 and 5.46). Therefore, the EM equations (until third order) for a multivariate GCSMM parameter estimation can be expressed as:

$$\begin{aligned}
\alpha^{(g)} = \frac{1}{N} \sum_j \tau_j^{(g)}, \boldsymbol{\mu}^{(g)} = \frac{1}{N} \sum_j \frac{\tau_j^{(g)} \mathbf{x}_j}{\alpha^{(g)}}, \mathbf{P}^{(g)} = \frac{1}{N} \sum_j \frac{\tau_j^{(g)} (\mathbf{x}_j - \boldsymbol{\mu}^{(g)}) (\mathbf{x}_j - \boldsymbol{\mu}^{(g)})^T}{\alpha^{(g)}} \\
\mathbf{P}^{(3)(g)} = \frac{1}{N} \sum_j \frac{\tau_j^{(g)} (\mathbf{x}_j - \boldsymbol{\mu}^{(g)}) (\mathbf{x}_j - \boldsymbol{\mu}^{(g)})^T \otimes (\mathbf{x}_j - \boldsymbol{\mu}^{(g)})^T}{\alpha^{(g)}} \tag{5.48}
\end{aligned}$$

where, j is the subscript for j^{th} data vector and N is the total number of data vectors. Computationally more involved higher order multivariate EM moment estimates are:

$$\begin{aligned}
\mathbf{P}^{(4)(g)} = \frac{1}{N} \sum_j \frac{\tau_j^{(g)} (\mathbf{x}_j - \boldsymbol{\mu}^{(g)}) (\mathbf{x}_j - \boldsymbol{\mu}^{(g)})^T \otimes_1^2 (\mathbf{x}_j - \boldsymbol{\mu}^{(g)})^T}{\alpha^{(g)}} \\
\mathbf{P}^{(5)(g)} = \frac{1}{N} \sum_j \frac{\tau_j^{(g)} (\mathbf{x}_j - \boldsymbol{\mu}^{(g)}) (\mathbf{x}_j - \boldsymbol{\mu}^{(g)})^T \otimes_1^3 (\mathbf{x}_j - \boldsymbol{\mu}^{(g)})^T}{\alpha^{(g)}} \tag{5.49}
\end{aligned}$$

where, the time subscript “ k ” has been omitted for clarity and replaced with data vector variable “ j ” and \otimes_1^x denotes the Kronecker product from 1 to x times e.g., $\otimes_1^3 A = \otimes A \otimes A \otimes A$. Equations:

5.43 can now be used to convert moments into cumulants in order to fully parameterize the functional form of GCSMM given in Equation: 5.47. Usually these Kronecker products are not required to be implemented in a computer programme. Instead only unique order moments are calculated using vectorized methods. This makes the computation faster and more efficient. The approximate *posterior* probability calculations in the EM algorithm are computed using each Gaussian component of GCSMM to avoid numerical instability. Numerical instability could arise due to likely negative *posterior* probabilities produced by the lower order GCS, which may result into negative weights $\alpha^{(g)}$ or negative diagonals of covariance matrices $\mathbf{\mu}^{(g)}$ (making them non-positive definite). Hence by using only Gaussian a component of each GCS, one can acquire positive *posterior* probabilities and always ensure avoidance of their probable negative regions. The computation of *posterior* probabilities is expressed as:

$$\tau_j^{(g)} \approx \frac{\alpha^{(g)} \mathcal{N}(\mathbf{x}_j, \mathbf{\mu}^{(g)}, \mathbf{P}^{(g)})}{\sum_{g=1}^G \alpha^{(g)} \mathcal{N}(\mathbf{x}_j, \mathbf{\mu}^{(g)}, \mathbf{P}^{(g)})} \quad (5.50)$$

where, $\tau_j^{(g)}$ is the estimated *posterior* probability that point \mathbf{x}_j is associated to the g^{th} main Gaussian term $\mathcal{N}(\mathbf{x}_j, \mathbf{\mu}^{(g)}, \mathbf{P}^{(g)})$ of Equation: 5.47.

Construction of GCS as corrections to a Gaussian PDF makes this a justified proposition. The EM algorithm is an iteration based algorithm thus, the parameters $\alpha^{(g)}$, $\mathbf{\mu}^{(g)}$, $\mathbf{P}^{(g)}$, $\mathbf{P}^{(3)(g)}$, $\mathbf{P}^{(4)(g)}$ and $\mathbf{P}^{(5)(g)}$ for each component of GCSMM as given in Equations: 5.48 and 5.49 need some initial values. In order to initialize the EM, we used the *k-means* algorithm of [109],[22]. The algorithm (*k-means*) is an essential tool for clustering of data in pattern recognition applications such as image analysis. Therefore, here the data vector \mathbf{x}_j (where, $j = 1 \dots N$ and N is the total number of data vectors) are partitioned into G clusters (where G is the number of clusters). Each cluster is represented by a mean vector $\mathbf{\mu}^{(g)}$ (where, $g = 1 \dots G$) and each data vector \mathbf{x}_j is assigned to a particular cluster based on its closest Euclidean distance vector to $\mathbf{\mu}^{(g)}$ expressed as [67]:

$$\sqrt{(\mathbf{x}_j - \mathbf{\mu}^{(g)})^T (\mathbf{x}_j - \mathbf{\mu}^{(g)})} \quad (5.51)$$

This algorithm also works iteratively wherein at each iteration the N data vectors are partitioned into G disjoint clusters S_g . An error function that is minimized is the total within the cluster sum of squares

expressed as [109]:

$$\mathfrak{E} = \sum_{g=1}^G \sum_{j \in S_g} \|\mathbf{x}_j - \boldsymbol{\mu}^{(g)}\|^2 \quad (5.52)$$

where, $\|\cdot\|$ is the Euclidean distance.

The initial partition of clusters is random. The condition expressed in Equation: 5.52 is checked at each iteration until further change in the error function is below a certain threshold. The initial parameters (statistics: $\boldsymbol{\mu}^{(g)}$, $\mathbf{P}^{(g)}$, $\mathbf{P}^{(3)(g)}$, $\mathbf{P}^{(4)(g)}$ and $\mathbf{P}^{(5)(g)}$ for each cluster) are then computed using data vectors in each cluster.

The initial parameters obtained from the output of *k-means* are provided to the EM algorithm. A likelihood function is defined in the EM algorithm which has to be maximized. Here a likelihood function is formulated, again using only the Gaussian component of GCSMM and defined as:

$$\Lambda(\mathbf{x}, \Xi) \approx \prod_{j=1}^N \sum_{g=1}^G \alpha^{(g)} \mathcal{N}(\mathbf{x}_j, \boldsymbol{\mu}^{(g)}, \mathbf{P}^{(g)}) \quad (5.53)$$

where, Ξ is a matrix with parameters of GCSMM.

The characterization of the maximum of likelihood function is done by using its logarithm [110]. Therefore, by taking the logarithm of Equation: 5.53 we get:

$$\mathfrak{G} = \sum_{j=1}^N \log \sum_{g=1}^G \alpha^{(g)} \mathcal{N}(\mathbf{x}_j, \boldsymbol{\mu}^{(g)}, \mathbf{P}^{(g)}) \quad (5.54)$$

For more details on the EM algorithm see references [107],[108],[110]. The expression in Equation: 5.54, is checked at each iteration until the change in that value decreases than a certain threshold or the number of preselected iterations end [110]. A modified Matlab function for EM based parameter estimation of GCSMM is termed as `gcsmmfit` function which finally provides following estimates:

$$\begin{aligned} \hat{\Xi} = & (\alpha_k^{(1)}, \boldsymbol{\mu}_k^{(1)}, \mathbf{P}_k^{(1)}, \kappa_{i,j,l}^{(1)}, \kappa_{i,j,l,m}^{(1)}, \kappa_{i,j,l,m,n}^{(1)} \dots \\ & \dots \alpha_k^{(g)}, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}, \kappa_{i,j,l}^{(g)}, \kappa_{i,j,l,m}^{(g)}, \kappa_{i,j,l,m,n}^{(g)}) \end{aligned} \quad (5.55)$$

where, $\hat{\Xi}$ is the optimal parameters estimate and the superscript denotes the GCSMM individual component index.

If one critically views the above hybrid scheme (*k-means* and EM,) it appears that for any given

samples of non-Gaussian distributed data vectors, one can obtain its multivariate GCSMM density estimates in a much simplified manner. The simplification is based on main Gaussian component (see Equation: 5.50 and 5.53). This is justifiable because each GCS component of GCSMM is essentially an extension / correction of the Gaussian PDF in terms of cumulant (function of higher order moments) based coefficients and Hermite polynomials. The use of Hermite polynomials is well suited for this problem due to its orthogonality with respect to the Gaussian weighting function (see Section: 5.3). The use of the main Gaussian component based approximation is also supported by practical issues associated with the PF. For example, these filters employ a limited number (i.e., 200-1000) of particles for computationally tractable algorithms e.g., if one is computing only Gaussian statistics (mean and covariance) from 200 particles would surely be neglecting the higher order moment structure (i.e., skewness, kurtosis etc), practically existing in particles. Therefore, there should be some methodology in PF (especially *parametric bootstrap* PF) which can capture these higher order structures such as by using GCS or GCSMM. As a consequence of use of GCS or GCSMM in PF algorithms one also needs generation of such particles which approximate well the original particles. Therefore next section describes a random number generation of GCS.

5.5 Random Number Generation

One of the most vital components of Gaussian PDF based PF algorithms is the normal random numbers generator. Therefore, to utilize GCS in PF, we have developed two different types of random number generator for GCS in Matlab named as `randngcs` and `coprangcs`.

5.5.1 GCS Random Number Generator using Acceptance Rejection

The method of Acceptance Rejection (AR) [67] has been used for generation of GCS distributed random numbers. In AR, firstly we select a PDF $p_{AR}(\mathbf{x}_k)$ from which it is simple to generate a random \mathbf{x}_k . This random vector \mathbf{x}_k will be considered as a random vector of actual PDF $p_{gcs}(\mathbf{x}_k)$ with probability proportional to $\frac{p_{gcs}(\mathbf{x}_k)}{p_{AR}(\mathbf{x}_k)}$. In order to do this we have to define a constant “ c ” so as to adjust the height of $p_{AR}(\mathbf{x}_k)$ to be always more than $p_{gcs}(\mathbf{x}_k)$:

$$\frac{p_{gcs}(\mathbf{x}_k)}{p_{AR}(\mathbf{x}_k)} \leq c \quad \forall \mathbf{x}_k \quad (5.56)$$

Actually the vectors are generated from $cp_{AR}(\mathbf{x}_k)$ and only accepted if they fall under the curve of the desired PDF $p_{gcs}(\mathbf{x}_k)$. Those vectors which are outside this curve are rejected. To achieve maximum efficiency, the number of rejected vectors should be minimal [67]. See Table: 5-2 for details of its algorithm.

The procedure for `randngcs` is outlined below:

1. Estimate the maximum height of $p_{gcs}(\mathbf{x}_k)$ as expressed in Equation: 5.40 using Matlab `fminsearch` (This function finds unconstrained minimum of a multivariable function using the derivative free method known as Nelder-Mead Algorithm [78]). This value is used to get “ c ”.
2. Select Gaussian PDF $p_{AR}(\mathbf{x}_k) = \mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k)$ such that Equation: 5.56 is satisfied.
3. Generate a random number / vector \mathbf{x}_k from $\mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k)$.
4. Generate a Uniform random number \mathcal{U} between 0 and 1.
5. If following condition holds:

$$\mathcal{U} \leq \frac{p_{gcs}(\mathbf{x}_k)}{cp_{AR}(\mathbf{x}_k)}$$

6. Accept \mathbf{x}_k as the value from $p_{gcs}(\mathbf{x}_k)$ otherwise go to step 3.

Table 5-2: Description of AR algorithm for generation of GCS distributed random vectors \mathbf{x}_k

In order to evaluate the usefulness of `randngcs`, and in particular the ability of GCS to model non-Gaussian PDFs, we selected phase space distribution of a simple nonlinear pendulum. The simple pendulum nonlinear dynamics are of considerable interest to researchers due to its simple form [111]. The equations for un-damped dynamics are expressed as:

$$\ddot{\theta} = -\frac{g}{l} \sin \theta \quad (5.57)$$

where, $g = 9.8 \text{ m/s}^2$, l is the length of the pendulum string and θ is the angle in radians. A simple pendulum with time period $T \triangleq 2\pi\sqrt{l/g} = 1 \text{ sec}$ is selected. One may now proceed by providing approximately 10000 normally (Gaussian) distributed initial conditions of the angular position (θ) and angular velocity ($\dot{\theta}$) at time = 0 sec to this pendulum. The collection of final conditions (or particles) of angular position and angular velocity after time = 10 sec is now considered for PDF estimation. Firstly, a non-parametric based PDF estimation result is shown in Figure: 5-6. The black contour lines on top represents the multivariate “Gaussian kernel \mathbf{K} ” based non-parametric density estimation of these particles. The equation for this PDF approximation is expressed as [67]:

$$p_{Ker}(\mathbf{x}_k) = \frac{1}{N\tilde{h}_1 \dots \tilde{h}_d} \sum_{i=1}^N \left\{ \prod_{j=1}^d \mathbf{K} \left(\frac{x_j - X_{ij}}{\tilde{h}_j} \right) \right\} \quad (5.58)$$

where, X_{ij} is the j^{th} component of i^{th} data particle and \tilde{h}_j is the smoothing parameter (usually a function of moments of the distribution).

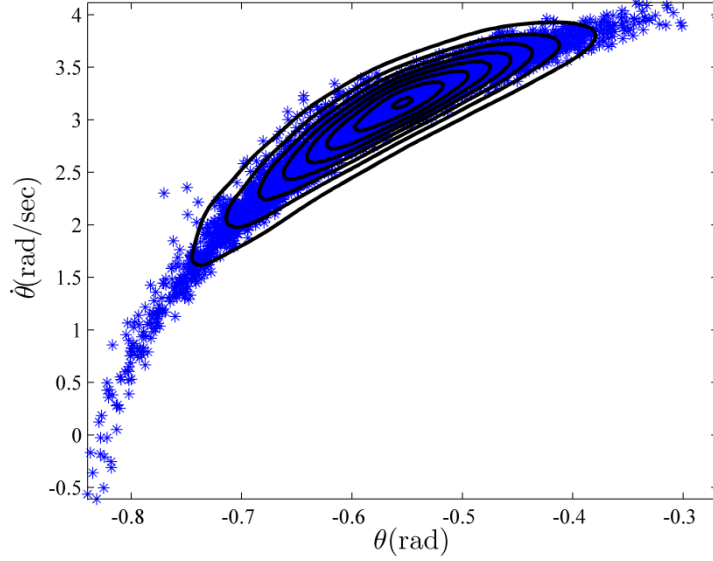


Figure 5-6: Gaussian kernel based non-parametric density estimation for simple nonlinear pendulum. The black lines show PDF contours over its true final conditions (particles).

Now consider a single Gaussian, GCS (5th order), GMM and GCSMM types, for PDF estimation of final conditions of the simple pendulum. The results are shown in Figures: 5-7 to 5-10. A single Gaussian distribution is clearly incapable of capturing the bent (skewness) and layout of particles which produces lot of gaps in estimated PDF (Figure: 5-7); whereas, GCS approximations appear to be more close fitting distribution (Figure: 5-8). However, note GCS does not estimate well near centroid of the PDF. Moreover, it is unable to capture the skewness (extended tails) of a PDF. GCSMM and GMM estimation results are closely related, one may relate the similarities in their structure and form (see Figure: 5-9 and 5-10). These estimations are based on three mixture components each, which appear to provide sufficient accuracy.

The random number generation for different PDFs using a particular type of algorithm vary in the output and may produce changing levels of noise. The results shown in the comparisons (Figure: 5-7 to 5-10) are basically estimation of outputs of Matlab built in `randn` and our `randngcs`, using a

Gaussian kernel estimator (Equation: 5.58). This serves as an independent test or criteria for judging the efficiency of a particular random number generator for use in any PF algorithm.

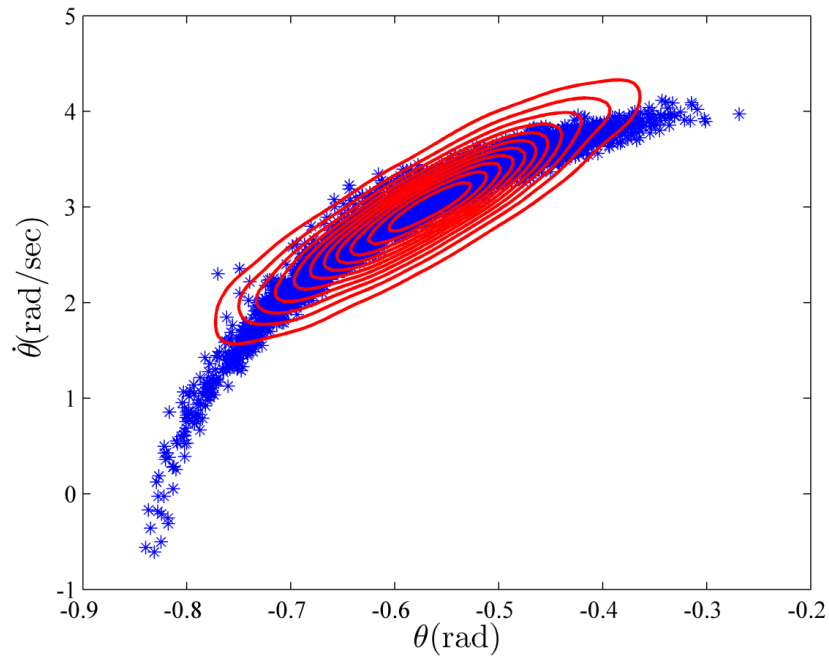


Figure 5-7: Single Gaussian PDF contours (red) for a Matlab (built in) “randn” generator plotted over true final conditions (particles in blue).

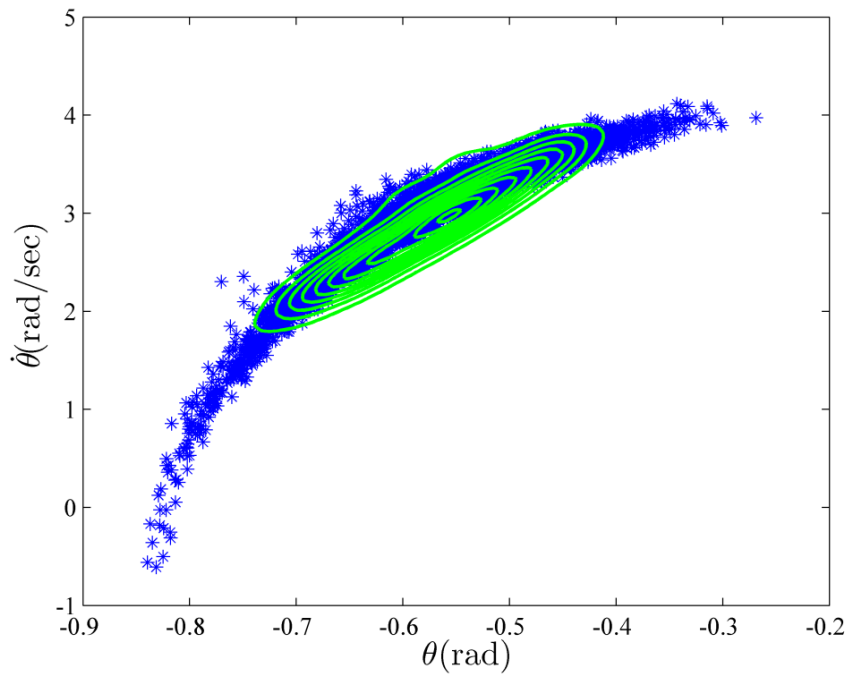


Figure 5-8: Single GCS (5th order) PDF contours (green) for a Matlab “randngcs” generator over true final conditions (particles in blue).

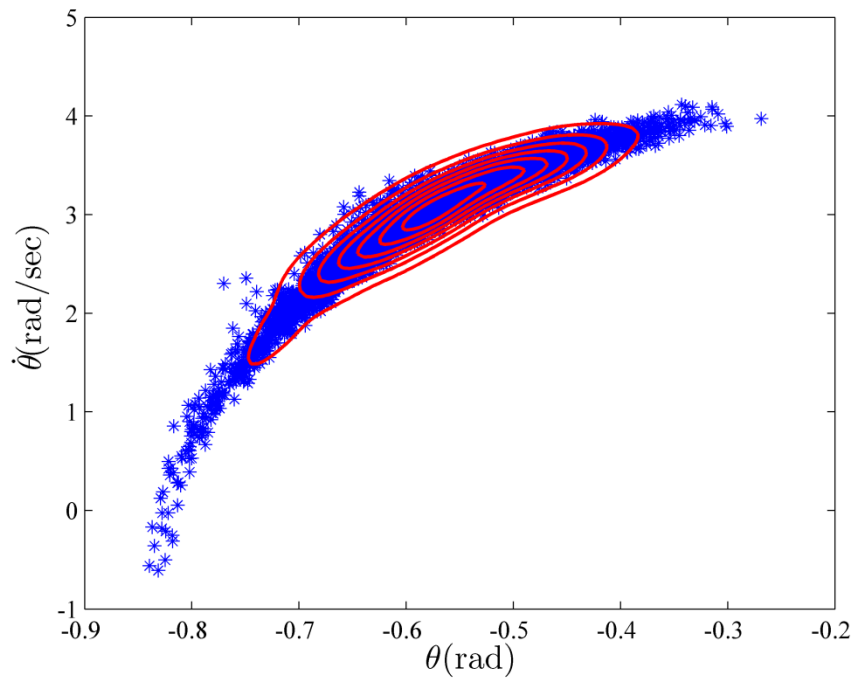


Figure 5-9: Three components GMM PDF contours (red) over true final conditions (particles in blue). Each component of GMM is generated using Matlab (built in) “randn” generator.

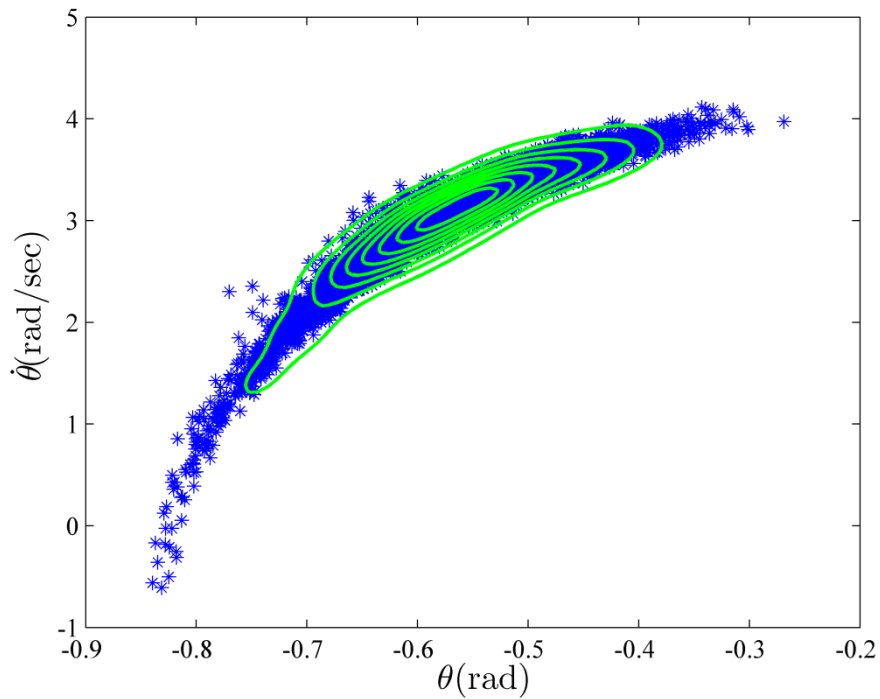


Figure 5-10: Three components GCSMM (5th order) PDF contours (green) over true final conditions (particles in blue). Each component of GCSMM is generated using Matlab “randngcs” generator.

5.5.2 Gram Charlier Series Random Number Generator using Gaussian Copula

Another random number generator for GCS in Matlab is also developed and is named as `coprاندngcs` based on Gaussian *copula*. The word “*copula*” is a Latin word for “bond” or “link”. It is a function which defines dependencies among variables and is used to generate correlated multivariate random numbers with specific marginal PDFs [112]. Given a joint PDF of two random variables, a marginal PDF of one variable is obtained by integrating the joint PDF over the other variable [5]. Thus, one may now express a bivariate Gaussian *copula* function as [112]:

$$C(u, v) = \mathcal{N}_\rho(\phi^{-1}(u), \phi^{-1}(v)) \quad (5.59)$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} e^{-\left(\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right)} ds dt$$

where, u and v are the marginal CDFs for bivariate random numbers and ρ is the correlation coefficient. In `coprاندngcs` GCS is chosen as marginal PDFs and use ρ from Gaussian PDF to generate correlations. Table: 5-3 describes the algorithm which is used to generate Gaussian *copula* based random numbers with GCS marginals up to order three, which can be extended to higher orders in similar manner. The Gaussian *copula* method is computationally more attractive than AR as it, (1) avoids the computation of higher order cross moments, and (2) aptly incorporates correlations between variables using rank correlation or linear correlation parameter. The effectiveness of `coprاندngcs` based PF over Gaussian or GMM based PF would be presented in the later part of the chapter. Again comparison on the lines of Section: 5.5.1 for `coprاندngcs` is carried out. The result is shown in Figure: 5.11 which is comparable to the results illustrated in Figure: 5-10.

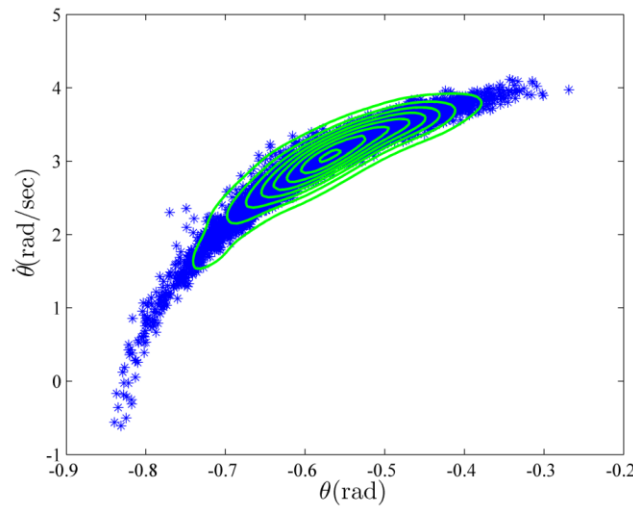


Figure 5-11: Three components GCSMM (3rd order) PDF contours (green) over true final conditions (particles in blue). Each component of GCSMM is generated using “`coprاندngcs`” generator.

1. Consider mean μ , variance σ^2 , skew μ_3 are available for each dimension separately.
2. Compute linear correlation ρ or (rank correlation [26]) to construct dependency. For example the multivariate linear correlation is expressed as:

$$\rho_{ij} = \frac{P_{ij}}{\sqrt{P_{ii}P_{jj}}}$$

where,

P_{ij} , ij component of covariance matrix \mathbf{P}

3. Establish grid for each dimension.
4. Convert statistics from step (1) to standardized cumulants $\kappa_i = \frac{\mu_i}{\sigma^i}$ where κ_i are cumulants expressed in terms of central moments (only first four are shown):

$$\kappa_2 = \sigma^2 = \mu_2$$

$$\kappa_3 = \mu_3$$

$$\kappa_4 = \mu_4 - 3\mu_2^2$$

5. Compute CDFs for GCS marginals for each dimension as:

$$\text{CDF} = \int_{-\infty}^{x_i} p_G(t_i, \mu, \sigma) \left[1 + \frac{1}{3!} \kappa_3 h_3(t_i, \mu, \sigma) + \dots \right] dt_i$$

6. Compute inverse CDF $\phi^{-1}(u_i)$ for each dimension from step (5) by inverting the function.
7. Generate Gaussian *copula* random variables with dependency structure (correlations) as in step (2) using Matlab `copularnd` function the function generates correlated multivariate Gaussian random variables domain (-1, 1).
8. Generate vector random variables \mathbf{x}_k by table look up method of probabilities from step (3) and (6) with correlations structure provided by step (7).

Table 5-3: Description of the algorithm for Gaussian *copula* based random number generator for GCS.

5.6 Gram Charlier Series and its Mixture Particle Filtering

This section describes the new estimation algorithms that are developed for estimation of nonlinear dynamical systems, such as estimating satellite orbits. As presented in Section: 5.3 to 5.4, GCS is found to be a natural extension of a Gaussian PDF. The higher order moments of any PDF can be aptly incorporated into the GCS or GCSMM formulation (Equation: 5.40 and 5.47), albeit making it a complex proposition. Now the basic ideas presented in generic (SIS-R) PF, GPF and GSPF algorithms

is extended by relaxing the Gaussian assumption with more useful GCS and GCSMM. The basic concept of PF remains the same with some dissimilarity among different algorithmic schemes. The following algorithms are developed and implemented: (1) Single GCS Particle Filter (GCSPF), (2) GCS Mixture Particle Filter (GCSMPF), and (3) Hybrid GCS Culver Particle Filter (HGCPF). As far as the author is aware use of GCS or GCSMM type of PDF in SMC filtering has not been previously attempted or considered in the literature. Moreover, the restriction on use of GCS for PF is also removed by fast generation of random numbers (see Table: 5-4).

Random generator	Number of bivariate random vectors generated	Time (sec)	Order of Hermite polynomial
randngcs	200	0.17	5 th
coprangcs	250	0.09	3 rd

Table 5-4: Random number generation timings for `randngcs` and `coprangcs` for order 5 and 3, respectively.

5.6.1 Single Gram Charlier Series Particle Filtering

The single GCS (Equation: 5.40) can be conveniently used for the SMC Bayesian filtering of nonlinear dynamical systems. The basic idea of the GCSPF is adapted from GPF. However, in GCSPF one approximates the Bayes' *a posteriori* PDF by GCS. GCS is considered as point mass approximated PDF. The function `randngcs` will be employed for generation of random vectors. In order to present algorithms for PF we proceed with GCS up to order five of Hermite polynomials. The extension to higher orders is possible but it would be computationally very expensive (see Figure: 1-4 and Section: 1.2). Compact notation of GCS up to order five is expressed as:

$$p_{gcs}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k, \mathbf{P}_k^{(3)}, \mathbf{P}_k^{(4)}, \mathbf{P}_k^{(5)}) \quad (5.60)$$

where, \mathbf{x}_k denotes the state for multivariate statistics $\boldsymbol{\mu}_k$ denotes a mean, \mathbf{P}_k is the covariance matrix, $\mathbf{P}_k^{(3)}$ is the coskewness tensor, $\mathbf{P}_k^{(4)}$ is the fourth order tensor, and $\mathbf{P}_k^{(5)}$ is the fifth order tensor. In this compact notation (Equation: 5.60) and more to follow, moments are used instead of cumulants (because of the convenient conversion relations between moments into cumulants given in Equation:

5.43) for the sake of simplicity. Two different algorithms GCSPF and HGCPF for SMC filtering employing truncated GCS up to order five and three respectively will be described. The former considers noise PDF also as GCS, expressed compactly as:

$$p_{gcs}(\mathbf{w}_k, \boldsymbol{\mu}_{\mathbf{w},k}, \mathbf{Q}_k, \mathbf{Q}_k^{(3)}, \mathbf{Q}_k^{(4)}, \mathbf{Q}_k^{(5)}) \quad (5.61)$$

where, the notation for noise variable is \mathbf{w}_k . The statistics for this variable $\boldsymbol{\mu}_{\mathbf{w},k}$ is the mean, \mathbf{Q}_k , is the covariance $\mathbf{Q}_k^{(3)}$ is the coskewness tensor, $\mathbf{Q}_k^{(4)}$ is the fourth order tensor, and $\mathbf{Q}_k^{(5)}$ is the fifth order tensor. Using Equation: 5.6 and the property of delta function [104] the state transition PDF can be expressed as:

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{x}_{k-1}) &= \int_{-\infty}^{+\infty} \delta(\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1})) p_{gcs}(\mathbf{w}_k, \boldsymbol{\mu}_{\mathbf{w},k}, \mathbf{Q}_k, \mathbf{Q}_k^{(3)}, \mathbf{Q}_k^{(4)}, \mathbf{Q}_k^{(5)}) d\mathbf{w}_k \\ &= p_{gcs}(\mathbf{x}_k, \mathbf{f}(\mathbf{x}_{k-1}) + \boldsymbol{\mu}_{\mathbf{w},k}, \mathbf{Q}_k, \mathbf{Q}_k^{(3)}, \mathbf{Q}_k^{(4)}, \mathbf{Q}_k^{(5)}) \end{aligned} \quad (5.62)$$

Now by using Equation: 5.5, the state predictive PDF can be derived as:

$$\begin{aligned} p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{k-1}) &= \int_{-\infty}^{+\infty} p_{gcs}(\mathbf{x}_k, \mathbf{f}(\mathbf{x}_{k-1}) + \boldsymbol{\mu}_{\mathbf{w},k}, \mathbf{Q}_k, \mathbf{Q}_k^{(3)}, \mathbf{Q}_k^{(4)}, \mathbf{Q}_k^{(5)}) \\ &\quad \times p_{gcs}(\mathbf{x}_{k-1}, \boldsymbol{\mu}_{k-1}, \mathbf{P}_{k-1}, \mathbf{P}_{k-1}^{(3)}, \mathbf{P}_{k-1}^{(4)}, \mathbf{P}_{k-1}^{(5)}) d\mathbf{x}_{k-1} \\ p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{k-1}) &= E \left[p_{gcs}(\mathbf{x}_k, \mathbf{f}(\mathbf{x}_{k-1}) + \boldsymbol{\mu}_{\mathbf{w},k}, \mathbf{Q}_k, \mathbf{Q}_k^{(3)}, \mathbf{Q}_k^{(4)}, \mathbf{Q}_k^{(5)}) \right] \end{aligned} \quad (5.63)$$

where, the expectation operator $E[\cdot]$ or marginalization in second of Equation: 5.63 is performed using the Bayes' *a posteriori* PDF at $(k-1)$ instant. One may consider a solution of Equation: 5.63 as GCS, therefore the statistics of samples i.e., mean, covariance using Equation: 5.30, and higher order moment tensors of state predictive PDF at k^{th} instant are expressed as:

$$\begin{aligned} \mathbf{P}_k^{(3)} &= \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k) (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \otimes (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \\ \mathbf{P}_k^{(4)} &= \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k) (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \otimes_1^2 (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \\ \mathbf{P}_k^{(5)} &= \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k) (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \otimes_1^3 (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \end{aligned} \quad (5.64)$$

Now consider the state predictive PDF as a *proposal* PDF. Therefore, we now substitute the *proposal*

PDF from Equation: 5.63 in Equation: 5.20. Thus, in the *measurement update* step, the evaluation of weights is computed in a non-iterative manner using $w_k^{(i)} = p(\mathbf{y}_k | \mathbf{x}_k = \mathbf{x}_k^{(i)})$ where, the samples $\mathbf{x}_k^{(i)}$ are drawn from the state predictive PDF during the *time update* step. The Bayesian inference is then drawn by computing weighted statistics of the Bayes' *a posteriori* PDF at " k^{th} " step. A pseudo-code for the algorithm pertaining to fifth order GCSPF is described in Table: 5-5. In HGCPF third order GCS is used to approximate the Bayes' *a posteriori* PDF. The word hybrid is used in this filter to signify use of third order CF *measurement update* equations. CF has already been briefly described in Chapter: 1 and 2. It pertains to *continuous-discrete* type of filtering therefore we would describe it in more detail, later in Chapter: 6. However, a brief description of the filter is given in this section as well. In the CF, a second order Taylor series linearization of the dynamical system is carried out in order to express the higher order moment (until third order) evolution equations of the system using the Ito differential rule [1],[6]. During the *time update*, these evolution equations are integrated forward in time to obtain parameters (moments) of the state predictive PDF approximated as GCS. Much like, the EKF the measurement function is also linearized to obtain the measurement likelihood also as Gaussian. MMSE and higher order moment solutions are found by analytically solving Bayes' formula. The measurement update equations for the CF are expressed as [1]:

$$\mu_i^+ = \mu_i + d_i + \phi_i \quad (5.65)$$

$$P_{ij}^+ = \Omega_{ij} + \psi_{ij} - \phi_i \phi_j$$

$$P_{ijl}^{(3)+} = N_{ijl} - \phi_i \psi_{jl} - \phi_j \psi_{il} - \phi_l \psi_{ij} + 2\phi_i \phi_j \phi_l$$

where, μ_i^+ is the mean of Bayes' *a posteriori* PDF for dimension (i), P_{ij}^+ is the covariance of Bayes' *a posteriori* PDF between dimensions (ij), $P_{ijl}^{(3)+}$ is the coskewness of Bayes' *a posteriori* PDF between dimension (ijl)

The other variables expressed on right hand side in Equation: 5.65 are described in Chapter: 6. In HGCPF time update equations, instead of linearizing the function using Taylor series the propagation of nonlinear function (first of Equation: 5.1) is carried out without any linearization. However, the measurement function is linearized as in original CF. The process noise in HGCPF is considered as additive Gaussian. However, a non-Gaussian process noise can also be considered. The state predictive PDF in HGCPF is approximated as done in GCSPF where the sample statistics of this PDF are computed on the lines of Equation: 5.30 and 5.64. See Table: 5-6 for the pseudo-code of HGCPF.

From time $k-1$...,

Time Update

For each $i = 1, \dots, N$, obtain samples

$$\{\mathbf{x}_{(k-1)}^{(i)}\}_{i=1}^N \sim p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) = p_{gcs}(\mathbf{x}_{k-1}, \boldsymbol{\mu}_{k-1}, \mathbf{P}_{k-1}, \mathbf{P}_{k-1}^{(3)}, \mathbf{P}_{k-1}^{(4)}, \mathbf{P}_{k-1}^{(5)})$$

For each $i = 1, \dots, N$, obtain samples

$$\{\mathbf{x}_{(k)}^{(i)}\}_{i=1}^N \sim p_{gcs}(\mathbf{x}_k; \mathbf{f}(\mathbf{x}_{k-1} = \mathbf{x}_{(k-1)}^{(i)}) + \boldsymbol{\mu}_{\mathbf{w}_{k-1}}, \mathbf{Q}_{k-1}, \mathbf{Q}_{k-1}^{(3)}, \mathbf{Q}_{k-1}^{(4)}, \mathbf{Q}_{k-1}^{(5)})$$

For $\{\mathbf{x}_k^{(i)}\}_{i=1}^N$ are distributed as GCS samples, obtain mean $\boldsymbol{\mu}_k$, covariance $\mathbf{P}_k, \mathbf{P}_k^{(3)}, \mathbf{P}_k^{(4)}$, and $\mathbf{P}_k^{(5)}$:

$$\boldsymbol{\mu}_k = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_k^{(i)}, \mathbf{P}_k = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T$$

$$\mathbf{P}_k^{(3)} = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \otimes (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T, \text{ and } \mathbf{P}_k^{(4)*}, \mathbf{P}_k^{(5)**}$$

Measurement Update

For $i = 1, \dots, N$, obtain samples from

$$\{\mathbf{x}_k^{(i)}\}_{i=1}^N \sim p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) = p_{gcs}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k, \mathbf{P}_k^{(3)}, \mathbf{P}_k^{(4)}, \mathbf{P}_k^{(5)})$$

For each $i = 1, \dots, N$, compute weights $w_k^{(i)} = p(\mathbf{y}_k | \mathbf{x}_k = \mathbf{x}_k^{(i)})$, $w_k^{(i)} = \frac{w_k^{(i)}}{\sum_{j=1}^N w_k^{(j)}}$

Weighted statistics /Inference from filtering density $\boldsymbol{\mu}_k, \mathbf{P}_k, \mathbf{P}_k^{(3)}, \mathbf{P}_k^{(4)}, \mathbf{P}_k^{(5)}$ are:

$$\boldsymbol{\mu}_k = \sum_{i=1}^N w_k^{(i)} \mathbf{x}_k^{(i)}, \mathbf{P}_k = \sum_{i=1}^N w_k^{(i)} (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T$$

$$\mathbf{P}_k^{(3)} = \sum_{i=1}^N w_k^{(i)} (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \otimes (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T$$

$$\mathbf{P}_k^{(4)} = \sum_{i=1}^N w_k^{(i)} (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \otimes_1^2 (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T$$

$$\mathbf{P}_k^{(5)} = \sum_{i=1}^N w_k^{(i)} (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \otimes_1^3 (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T$$

*,** See Equation: 5.64

Table 5-5: The Gram Charlier Series Particle Filter (GCSPF)

From time $k-1$...

Time Update

For each $i = 1, \dots, N$, obtain samples

$$\{\mathbf{x}_{(k-1)}^{(i)}\}_{i=1}^N \sim p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) = p_{gcs}(\mathbf{x}_{k-1}, \boldsymbol{\mu}_{k-1}, \mathbf{P}_{k-1}, \mathbf{P}_{k-1}^{(3)})$$

For each $i = 1, \dots, N$, obtain samples

$$\{\mathbf{x}_{(k)}^{(i)}\}_{i=1}^N \sim p_{gcs}(\mathbf{x}_k; \mathbf{f}(\mathbf{x}_{k-1} = \mathbf{x}_{(k-1)}^{(i)}) + \boldsymbol{\mu}_{\mathbf{w}_{k-1}}, \mathbf{Q}_{k-1})$$

For $\{\mathbf{x}_k^{(i)}\}_{i=1}^N$ are distributed as GCS samples, obtain mean $\boldsymbol{\mu}_k$, covariance $\mathbf{P}_k, \mathbf{P}_k^{(3)}$:

$$\boldsymbol{\mu}_k = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_k^{(i)}, \mathbf{P}_k = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T$$

$$\mathbf{P}_k^{(3)} = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)(\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T \otimes (\mathbf{x}_k^{(i)} - \boldsymbol{\mu}_k)^T$$

Measurement Update

Use Culver Filter (CF) measurement update equations [1](See details for CF in Chapter: 6)

$$\boldsymbol{\mu}_i^+ = \boldsymbol{\mu}_i + \mathbf{d}_i + \boldsymbol{\phi}_i$$

$$\mathbf{P}_{ij}^+ = \Omega_{ij} + \psi_{ij} - \phi_i \phi_j$$

$$\mathbf{P}_{ijl}^{(3)+} = \mathbf{N}_{ijl} - \phi_i \psi_{jl} - \phi_j \psi_{il} - \phi_l \psi_{ij} + 2\phi_i \phi_j \phi_l$$

Measurement updated Bayes' a posteriori PDF

$$p_{gcs}(\mathbf{x}_k, \boldsymbol{\mu}_k, \mathbf{P}_k, \mathbf{P}_k^{(3)})$$

where,

$$\boldsymbol{\mu}_i^+ \in \boldsymbol{\mu}_k, \mathbf{P}_{ij}^+ \in \mathbf{P}_k, \mathbf{P}_{ijl}^{(3)+} \in \mathbf{P}_k^{(3)}$$

Table 5-6: The Hybrid GCS and Culver Particle PF (HGCPF)

5.6.2 Gram Charlier Series Mixture Particle Filtering

The improved fidelity of GCSMM has already been demonstrated in Section: 5.4 and 5.5. Therefore the truncated GCS up to order three in a mixture model configuration is used in nonlinear SMC filtering. Note that the order of GCS in this filter can be conveniently extended to higher orders using the same methodology. The Bayes' *a posteriori* of the state and noise PDF in this filter are considered as GCSMM. However, one may consider additive Gaussian noise also. The compact form of GCSMM can be expressed as:

$$p_{gcsm}(\mathbf{x}_k) = \sum_{g=1}^G \alpha_k^{(g)} p_{gcs}(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}, \mathbf{P}_k^{(3)(g)}) \quad (5.66)$$

An important point to note is the ability of the GCSMPF to incorporate (additive) highly non-Gaussian process noise expressed compactly as:

$$p_{gcsm}(\mathbf{w}_k) = \sum_{i=1}^I \beta_k^{(i)} p_{gcs}(\mathbf{w}_k, \boldsymbol{\mu}_{\mathbf{w},k}^{(i)}, \mathbf{Q}_k^{(i)}, \mathbf{Q}_k^{(3)(i)}) \quad (5.67)$$

During the *time update*, firstly the samples from the PDF expressed in Equations: 5.66 and 5.67 are drawn as per weights $\alpha_k^{(g)}$ and $\beta_k^{(i)}$. For example one may use the SIR as explained in Section: 5.2.4. These samples are propagated through the nonlinear dynamical system $\mathbf{f}(\cdot)$ (Equation: 5.1) just like SIS-R PF. By approximating the propagated distribution as GCSMM one employs an EM step using `gcsmmfit` function (see Section: 5.4.2) to obtain time updated “G” component state predictive GCSMM PDF. The *proposal* PDF in this filter is also considered as state predictive PDF available from the *time update*. Therefore, in the *measurement update* step the samples are redrawn from state predictive GCSMM PDF and the weights for “M” particles of each mixand (component) are computed using the observation likelihood $p(\mathbf{y}_k | \mathbf{x}_k = \mathbf{x}_k^{(i)})$ just as in the GSPF. Here Equation: 5.20 is used again. The weighted updates of parameters for each mixand are computed as:

$$\begin{aligned} \boldsymbol{\mu}_k^{(g)} &= \frac{\sum_{j=1}^M w_k^{(j)(g)} \mathbf{x}_k^{(j)(g)}}{\sum_{j=1}^M w_k^{(j)(g)}}, \mathbf{P}_k^{(g)} = \frac{\sum_{j=1}^M w_k^{(j)(g)} (\mathbf{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)}) (\mathbf{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)})^T}{\sum_{j=1}^M w_k^{(j)(g)}} \\ \mathbf{P}_k^{(3)(g)} &= \frac{\sum_{j=1}^M w_k^{(j)(g)} (\mathbf{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)}) (\mathbf{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)})^T \otimes (\mathbf{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)})^T}{\sum_{j=1}^M w_k^{(j)(g)}} \end{aligned} \quad (5.68)$$

The inference can now be conveniently drawn through parameters of GCSMM given in Equation: 5.68. The Pseudo-code for the filter is presented in Table: 5-7.

From time $k-1$...

Time Update

For $g = 1, \dots, G$, obtain samples as per the weights $\alpha_{k-1}^{(g)}$

$$\{\mathbf{x}_{(k-1)}^{(j)}\}_{j=1}^M \sim p_{gcsm}(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) = \sum_{g=1}^G \alpha_{k-1}^{(g)} p_{gcs}(\mathbf{x}_{k-1}, \boldsymbol{\mu}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{(g)}, \mathbf{P}_{k-1}^{(3)(g)})$$

For $i = 1, \dots, I$, obtain samples from $p_{gcsm}(\mathbf{w}_{k-1})$ as per weights $\beta_{k-1}^{(i)}$ and propagate through nonlinear system (first of Equation: 5.1).

Perform EM (`gcsmmfit`) step on propagated $\{\mathbf{x}_k^{(j)}\}_{j=1}^M$ particles to extract “ G ” component GCSMM time updated predictive PDF:

$$p_{gcsm}(\mathbf{x}_k | \mathbf{y}_{1:k-1}) = \sum_{g=1}^G \alpha_k^{(g)} p_{gcs}(\mathbf{x}_k, \boldsymbol{\mu}_k^{(g)}, \mathbf{P}_k^{(g)}, \mathbf{P}_k^{(3)(g)})$$

Measurement Update

For $g = 1, \dots, G$, obtain samples from $p_{gcs}^{(g)}(\mathbf{x}_k)$ and denote them as $\{\mathbf{x}_k^{(j)(g)}\}_{j=1}^M$.

For $g = 1, \dots, G$ each $j = 1, \dots, M$, compute weights, $w_k^{(j)(g)} = p(\mathbf{y}_k | \mathbf{x}_k = \mathbf{x}_k^{(j)(g)})$

For $g = 1, \dots, G$, Compute mean, covariance and tensor* components $\mathbf{P}_k^{(3)(g)}$:

$$\boldsymbol{\mu}_k^{(g)} = \frac{\sum_{j=1}^M w_k^{(j)(g)} \mathbf{x}_k^{(j)(g)}}{\sum_{j=1}^M w_k^{(j)(g)}}, \mathbf{P}_k^{(g)} = \frac{\sum_{j=1}^M w_k^{(j)(g)} (\mathbf{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)}) (\mathbf{x}_k^{(j)(g)} - \boldsymbol{\mu}_k^{(g)})^T}{\sum_{j=1}^M w_k^{(j)(g)}}$$

Update weights $\alpha_k^{(g)} = \alpha_{k-1}^{(g)} \frac{\sum_{j=1}^M w_k^{(j)(g)}}{\sum_g \sum_{j=1}^M w_k^{(j)(g)}}, \alpha_k^{(g)} = \frac{\alpha_k^{(g)}}{\sum_g \alpha_k^{(g)}}$

Inference: The conditional mean state estimate $\hat{\mathbf{x}}_k = E[\mathbf{x}_k | \mathbf{y}_{1:k}]$ and Covariance

$\hat{\mathbf{P}}_k = E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T]$ can be estimated by:

$$\hat{\mathbf{x}}_k = \sum_{g=1}^G \alpha_k^{(g)} \boldsymbol{\mu}_k^{(g)}, \hat{\mathbf{P}}_k = \sum_{g=1}^G \alpha_k^{(g)} \left(\mathbf{P}_k^{(g)} + (\boldsymbol{\mu}_k^{(g)} - \hat{\mathbf{x}}_k) (\boldsymbol{\mu}_k^{(g)} - \hat{\mathbf{x}}_k)^T \right)$$

Optional Step: Residual Resampling (Section: 5.2.4) applied on mixture weights to avoid use of insignificant (very small) weights in next time step.

*(See Equation: 5.68)

Table 5-7: The GCS Mixture PF (GCSMPF)

5.7 Experiments – Nonlinear Simple Pendulum

A nonlinear simple pendulum is standard example due to its interesting dynamical properties [111]. Particle representation of simple pendulum phase space in Figure: 5-6 depict non-Gaussian distributions. Therefore, state estimation of such a dynamical system under sparse measurement data requires better approximation of state predictive and Bayes' *a posteriori* PDF such as GCS or GCSMM.

5.7.1 Atmospheric Drag

The equations of motion of an undamped simple pendulum have already been expressed in Equation: 5.57. Here, we shall consider the damped simple pendulum's equation of motion. The damping accelerations are a suitable model for the effects of atmospheric or air drag accelerations upon a satellite orbit. The function used is a velocity-squared damping is [113]:

$$\ddot{\theta} = -\frac{g}{l} \sin \theta - 0.1211 \dot{\theta} |\dot{\theta}| \quad (5.69)$$

where, the constant = 0.1211 (coefficient of atmospheric drag acceleration), and the model for atmospheric drag upon satellites is expressed as [13],[12]:

$$\mathbf{a}_D = -\frac{1}{2} \rho \frac{C_D A}{m} v_r^2 \frac{\mathbf{v}_r}{|\mathbf{v}_r|} \quad (5.70)$$

where, $\mathbf{v}_r = |\mathbf{v}_r|$ (is the relative velocity of a satellite with respect to the atmosphere), m is the satellite mass, C_D is the drag coefficient (dimensionless quantity that describes the interaction of atmosphere with satellite's surface material), ρ is the atmospheric density at the location of satellite, and A is the satellite cross-sectional area. Also, like the un-damped pendulum this oscillator (Equation: 5.69) accounts for the nonlinear performance inherent in large amplitude swings [113]. Consider the model expressed in Equation: 5.69 as the *true model* for generating the *reference trajectory*. By a *reference trajectory* one means the state trajectory of the simple pendulum $(\theta(t), \dot{\theta}(t))$ that is being used to compare the output from the filters termed as *estimated trajectory*.

The discrete time measurement equation which gives the *reference trajectory* is described by:

$$y_k = \theta(t_k) + n_k \quad (5.71)$$

where, y_k is the k^{th} measurement of angular position $\theta(t_k)$ at the time instant t_k and n_k is the white Gaussian measurement noise due to sensor errors with the following statistics:

$$\begin{aligned} E[n_k] &= 0 \\ E[n_k n_l] &= R \delta_{kl} \end{aligned} \tag{5.72}$$

where, R is the correlation function and δ_{kl} is the Dirac delta function.

Consider next a *filter model*. Suppose that only the dynamics of an undamped simple pendulum is known and nothing is known about the damping acceleration ($0.1211\dot{\theta}|\dot{\theta}|$) as given in Equation: 5.69. The most appropriate choice to account for this unknown acceleration in a *filter model* is to formulate a model based on a Stochastic Differential Equation (SDE) [1]. In a SDE dynamics are described in terms of *deterministic* and *stochastic* forms. Therefore, a *filter model* is expressed for the simple pendulum using a SDE as [6],[23]:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}, t) + \mathbf{G}(\mathbf{x}, t)\mathbf{W}(t) \tag{5.73}$$

$$\mathbf{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}, \mathbf{f}(\mathbf{x}, t) = \begin{bmatrix} x_2 \\ -\frac{g}{l} \sin(x_1) \end{bmatrix}, \mathbf{G}(\mathbf{x}, t) = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_{22} \end{bmatrix},$$

where, $\mathbf{W}(t)$ is a white noise, $\mathbf{G}(\mathbf{x}, t)$ is the dispersion matrix and σ_{22} is the diffusion coefficient.

The white noise is considered as zero mean and its diffusion matrix is expressed as:

$$E[\mathbf{W}(t)\mathbf{W}(t + \tau)] = \mathbb{Q}(t)\delta(\tau) \tag{5.74}$$

Additive white noise inputs in Equation: 5.73 are based on the fact that the desired time correlation properties of a physically observed phenomena can be produced sufficiently well when white noise is passed through a linear shaping filter [23]. Thus, the term $\mathbf{G}(\mathbf{x}, t)\mathbf{W}(t)$ is augmented to basic *deterministic* dynamics to formulate a *stochastic* simple pendulum model. Another useful form of the SDE can be expressed as [23]:

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}, t)dt + \mathbf{G}(\mathbf{x}, t)d\boldsymbol{\beta}(t) \quad (5.75)$$

$$d\boldsymbol{\beta}(t) = \begin{bmatrix} 0 \\ d\beta \end{bmatrix}$$

where, $\boldsymbol{\beta}(t)$ is the Brownian motion vector, $d\beta$ is the scalar Brownian motion increment.

Here, we made use of the fact that, the white noise is derivative of the Brownian motion, $\mathbf{W}(t) = d\boldsymbol{\beta}(t)/dt$ [6],[23]. In general, the scalar Brownian motion over the time domain $[0, T]$ is defined as the random variable $\beta(t)$ that depends continuously on $t \in [0, T]$ and satisfies three conditions expressed in Table: 5-8 [114].

1. $\beta(0) = 0$ (with probability 1).
2. For $0 \leq t_{k-1} < t_k \leq T$ the random variable given by the increment $\beta(t_k) - \beta(t_{k-1})$ is normally distributed with mean = 0 and variance = $t_k - t_{k-1}$ or $\beta(t_k) - \beta(t_{k-1}) \approx \sqrt{t_k - t_{k-1}} \times \mathcal{N}(0,1)$, where $\mathcal{N}(0,1)$ = Gaussian distributed random variable with zero mean and unity variance.
3. For $0 \leq t_{k-1} < t_k < t_{k+1} < t_{k+2} \leq T$, the increments $\beta(t_k) - \beta(t_{k-1})$ and $\beta(t_{k+2}) - \beta(t_{k+1})$ are independent.

Table 5-8: The conditions defining the scalar Brownian motion process.

The formulation in Equation: 5.75 is basically an undamped simple pendulum dynamics (*deterministic* dynamics), added with a *stochastic* term $\mathbf{G}(\mathbf{x}, t)d\boldsymbol{\beta}(t)$ which accounts for unknown accelerations and state uncertainty as time progresses. Moreover, these dynamics are expressed in continuous time notation; therefore in order to utilize *discrete filtering* algorithms described earlier in this chapter, one has to express these into a *discrete time* formulation. In general, the solution of SDE (Equation: 5.73 and 5.75) is given by [23],[6] :

$$\mathbf{x}_k(t) = \mathbf{x}_{k-1}(t_0) + \int_{t_0}^t \mathbf{f}(\mathbf{x}_{k-1}(\tau), \tau)d\tau + \int_{t_0}^t \mathbf{G}(\mathbf{x}_{k-1}(\tau), \tau)d\boldsymbol{\beta}(\tau) \quad t \in (0, \Delta] \quad (5.76)$$

The first integral in Equation: 5.76 is an ordinary integral which can be solved usually through

numerical integration methods such as Runge-Kutta (RK-4). However, the evaluation of the second integral is not possible using ordinary differential calculus as the Brownian motion is a zero mean process with Gaussian increments, which is continuous but nowhere differentiable (with probability 1) [23],[6]. By considering the dispersion matrix $\mathbf{G}(\mathbf{x}, t)$ as diagonal (Equation: 5.73) one can approximate the solution of the second integral in Equation: 5.76 by considering Brownian motion as [114]:

$$\beta_k - \beta_{k-1} = d\beta(t) \approx \sqrt{\Delta t} \mathcal{N}(0,1) \quad (5.77)$$

Now one can select an appropriate fixed integration step size Δt for the solution of the first integral using RK-4 and solve the second integral of Equation: 5.76 using Equation: 5.77. This would yield an approximate discrete time solution of the SDE for a simple pendulum *filter model* (Equation: 5.73). In the conducted experiment the time period of the simple pendulum is selected as 1 sec. Therefore, a fixed step size is selected as $\Delta t = 0.01$ sec to solve both the *true model* and *filter model* in order to obtain high fidelity solution. In order to gauge the filtering performance one has to select some assessment criteria. The most direct approach is to evaluate the difference between the *true trajectory* and *estimated trajectory* at all instants of time. This would not only provide time history of Instantaneous Errors (IE), but also maximum and minimum error amplitudes, peculiarities in error pattern (i.e., periodicity / secular trends), convergence or divergence pattern and biases. However, another useful measure is to find out the standard deviation of errors over the complete estimated time span which provides us confidence in estimates. It could be found out by computing the square root of second moment of distribution of errors and termed as Root Mean Square Error (RMSE) or $1 \times$ standard deviation error. For example in case of Gaussian distributed errors, $1 \times$ standard deviation error in a particular dimension provides us 68.27% confidence that our errors at any time are within this value. Another important criterion of filtering performance is consistency in estimates. This could be measured by computing first moment (mean) of distribution of errors termed here as Mean Error (ME). Ideally, ME should be equal to zero for estimates to be termed as consistent [5]. Now we define the equations for these error criteria:

- IE

$$\delta \mathbf{x}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k \quad (5.78)$$

- RMSE

$$\sqrt{E[(\mathbf{x}_{0:T} - \hat{\mathbf{x}}_{0:T})^2]} \quad (5.79)$$

- ME

$$\delta \bar{\mathbf{x}} = E[\mathbf{x}_{0:T} - \hat{\mathbf{x}}_{0:T}] \quad (5.80)$$

where, $\mathbf{x}_{0:T}$ = *true* state from $k = 0$ to simulation time $k = T$, $\hat{\mathbf{x}}_{0:T}$ = *estimated* state from $k = 0$ to

simulation time $k = T$, and $E [.] =$ Expectation operator.

Section: 5.1 discussed the filtering requirements for nonlinear dynamical systems under less measurement availability pertaining to LEO space object radar orbit determination. Depending on the height of the orbit from surface of Earth a space object in LEO appears for a very short span of time (usually 5 to 10 minutes) on the horizon. This time span forms approximately $1/10^{\text{th}}$ of their orbital period. Moreover, even during the availability of satellite on the horizon does not necessarily guarantee useful measurements due to its attitude (angular orientation). The attitude is very important for useful / strong return of radar energy. Therefore, one may approximate the availability of measurements for only 2-3% of the total orbital period. Keeping in view, the above situation a 2% measurement availability time is selected for the simple pendulum problem. Since, the time period of simple pendulum is selected as 1 sec; therefore the measurements would only be available for 0.02 sec. This suggests only two measurements per time period in the experimental setup. The noise variance of a measurement sensor is a random parameter. For any radar system it can be estimated using MC simulations. Therefore, for the simulation purpose it is assumed as $R = 0.0017 \text{ rad}^2$ (Equation: 5.72) for all filters used in this experiment. For nonlinear dynamical systems under sparse (less) measurements the accuracy in the state predictive PDF becomes very crucial. As any optimal criteria the MMSE or MAP would then be acquired using this PDF. Table: 5-9 summarizes all the simulation parameters for our dynamic system including the initial conditions and initial error variances. The initial state uncertainty i.e., error covariance for satellites is usually large. Therefore, filtering under large initial state uncertainty is kept in mind while selecting error covariance for this experiment.

Model	Time Period	Integration step size (RK-4)	Initial Conditions		Simulation time
True	1 sec	0.01 sec	$\theta_0 = 0.7840 \text{ rad}$	$\dot{\theta}_0 = -0.2790 \text{ rad/sec}$	20 sec
Filter	1 sec	0.01 sec	$\theta_0 + \sqrt{P_{11}}$	$\dot{\theta}_0 + \sqrt{P_{22}}$	20 sec
where, P_{11} and P_{22} diagonal components of initial error covariance matrix (assumed diagonal) with following values (fixed for all the filters used in this experiment): $P_{11} = 0.01 \text{ rad}^2$ $P_{22} = 0.001 \text{ rad}^2(\text{sec}^{-2})$					

Table 5-9: Summary of parameters of dynamic system (simple pendulum atmospheric drag model) for simulation.

The number of particles in this experiment for PF family which includes GPF, GCSPF (using 5th order GCS) and generic (SIS-R) PF is selected as 200. Matlab Rebel toolkit [22] is applied to simulate filters (less algorithms based on GCS) in this experiment (the toolkit provides Bayesian *discrete time* sequential filters in generic form like EKF, PF and GSPF etc which are adapted for our simple pendulum experiment). The diffusion coefficient term σ_{22} expressed in Equation: 5.73, is the tuning parameter for optimal performance of the PF family (PF, GPF, and GCSPF) used in this experiment. See Figures: 5-12 and 5-13 for results of IE between the *true trajectory* and *estimated trajectory*. The time histories of errors in these figures are computed by averaging errors over 50 MC runs of each filter as suggested by ref [5] for comparison of different filtering algorithms. Note the single run of filter is considered to be a naive method for assessment of the performance [5]. The figures indicate that the performance of GCSPF is much better than other filtering algorithms. In GCSPF one can clearly identify more errors initially. However, the errors decrease and converge to relatively small amplitudes after 10 sec (approx). PF (SIS-R) performed reasonably well thereafter and has low initial errors and better comparative convergence. The statistics of ME in Table: 5-13 indicate better consistency for all the filters. The RMSE for GCSPF is slightly higher than the expected due to more errors during the first 10 sec of the simulation. However, having in mind its better convergence and low error amplitude subsequently; it can be considered comparatively better than other filters.

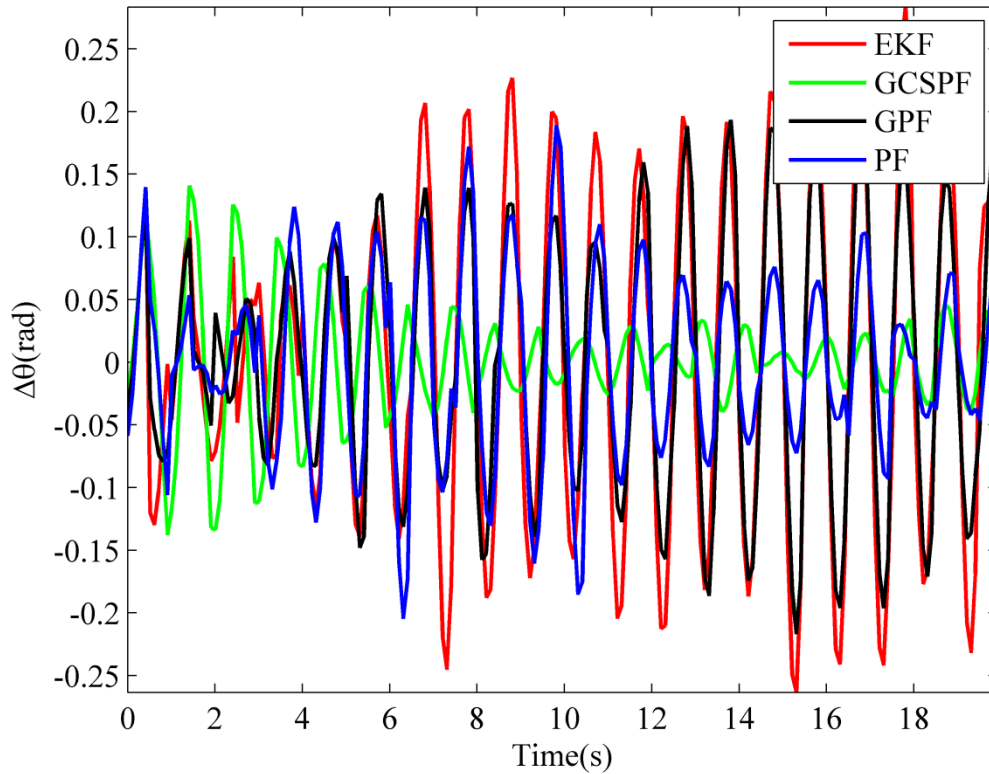


Figure 5-12: Comparison of time history of errors in angular position $\Delta\theta$ between true trajectory and estimated trajectory (atmospheric drag simple pendulum model) for different filters.

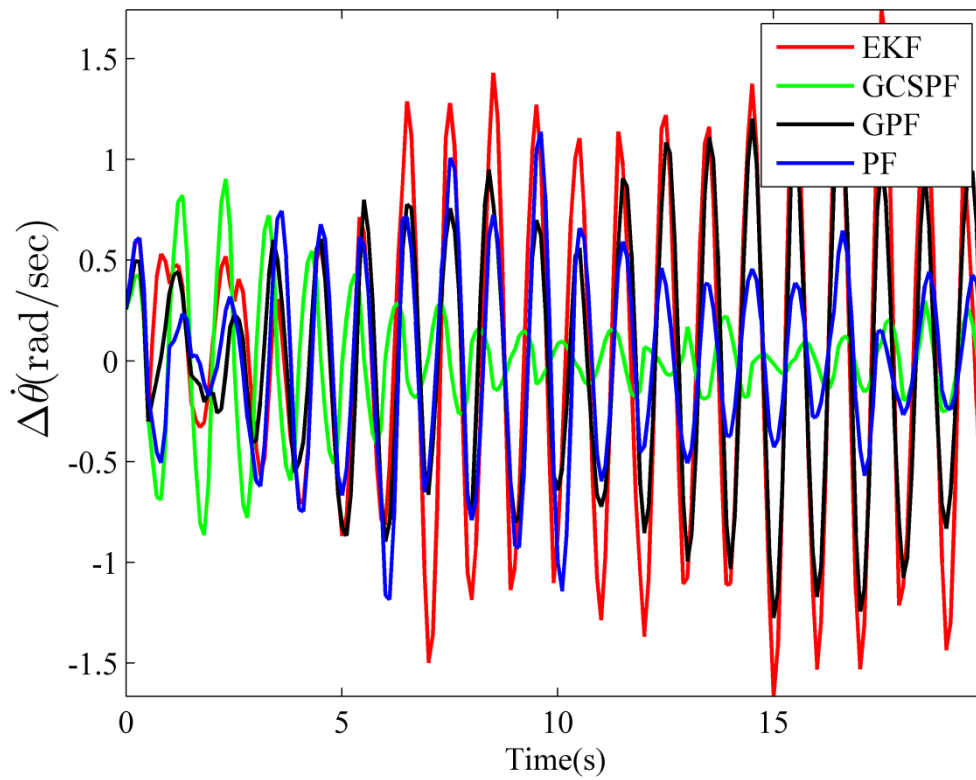


Figure 5-13: Comparison of time history of errors angular velocity $\Delta\dot{\theta}$ between true trajectory and estimated trajectory (atmospheric drag simple pendulum model) for different filters.

Filter	Mean Error (ME)		Root Mean Square Error (RMSE)	
	$\Delta\theta(\text{rad})$	$\Delta\dot{\theta}(\text{rad. sec}^{-1})$	$\Delta\theta(\text{rad})$	$\Delta\dot{\theta}(\text{rad. sec}^{-1})$
EKF	-0.0032	-0.0004	0.0987	0.6186
GCSPF	-0.0001	-0.0004	0.0945	0.5852
GPF	-0.0010	-0.0003	0.1733	1.0352
PF	-0.0008	-0.0008	0.0660	0.3941

Table 5-10: Comparison of filters in terms of ME and RMSE for atmospheric drag simple pendulum model.

5.7.2 Wind Gust

Aerospace launch vehicles which are used to transport satellites in specific orbits are imposed with structural loads due to wind variability along their trajectory through the Earth atmosphere into space. The most effective loads are due to discrete wind gusts which stand out above the general disturbance levels. The wind gusts are characterized by their length and amplitude which has Bivariate Gamma Distribution (BGD). The marginal PDF of this BGD are also univariate gamma PDF [115]. Therefore, it seems appropriate to add gamma distributed random numbers in nonlinear dynamics of the simple pendulum in order to simulate effects of wind gusts. We now describe *true model* for simple pendulum added with gamma distributed process noise:

$$\mathbf{x} = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}, \dot{\mathbf{x}} = \begin{bmatrix} x_2 \\ -\frac{g}{l} \sin(x_1) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \gamma \end{bmatrix} \quad (5.81)$$

The gamma PDF is given by [67]:

$$p(\gamma) = \frac{\lambda e^{-\lambda\gamma} (\lambda\gamma)^{t'-1}}{\Gamma(t')} \quad (5.82)$$

$$\Gamma(t') = \int_0^\infty e^{-y} y^{t'-1} dy$$

where, t' is the shape parameter, λ is the scale parameter and $\Gamma(t')$ is the gamma function. The *filter model* is same as expressed in the SDE Equation: 5.73. Table: 5-14 summarizes all the simulation parameters for our dynamic system including initial conditions and initial error variances.

Model	Time Period	Integration step size (RK-4)	Initial Conditions		Simulation time	Gamma PDF parameters*
True	1 sec	0.01 sec	$\theta_0 = 0.7840$ rad	$\dot{\theta}_0 = -0.2790$ rad.sec ⁻¹	20 sec	$t' = 0.27$ $\lambda = 0.27$
Filter	1 sec	0.01 sec	$\theta_0 + \sqrt{P_{11}}$	$\dot{\theta}_0 + \sqrt{P_{22}}$	20 sec	-
where, P_{11} and P_{22} diagonal components of initial error covariance matrix (assumed diagonal) with following values (fixed for all the filters used in this experiment): $P_{11} = 0.01$ rad ² $P_{22} = 0.001$ rad ² (sec ⁻²) *These parameters are fixed so as to obtain univariate gamma random numbers appropriate for simple pendulum used in this experiment ($\theta(t) \leq 0.7804$ rad).						

Table 5-11: Summary of parameters of dynamic system (simple pendulum wind gust model) for simulation.

The measurement model (Equation: 5.71), noise variances (Equation: 5.72) and number of measurements are kept the same as in the example for the atmospheric drag simple pendulum model in Section: 5.7.1. Now a comparison is performed for the same error criteria to evaluate filtering performance as already described in Equation: 5.78 to 5.80. Figures: 5-14 and 5-15 illustrate the IE between the *true* and *estimated trajectories* for the EKF, GCSPF, GSPF, (SIS-R) PF and HGCPF. One can clearly see sub optimality of performance of these filters compared with the results for the atmospheric drag model. This is mainly due to the additive gamma random numbers which are modelled as a white Gaussian noise in the *filter model* (see Equation: 5.73). However, even by using a SDE for the *filter model* one find, filters based on the GCS outperform others. Overall the performance of PF family (GCSPF, GSPF, and HGCPF excluding (SIS-R) PF) has shown an improvement over the EKF. The filter error statistics of ME and RMSE are shown in Table: 5-12 which shows comparatively better estimates can be achieved by using a filter based on GCS. However, the overall view of the RMSE and ME results suggests considerable angular deviation and inconsistency, respectively. The GSPF is initialized for two components (mixands) GMM; whereas, GCSPF is based on a single GCS which shows improvement over GMM. The number of particles in the PF family algorithms is selected as 200. This experiment is basically aimed at implementing the EKF, GSPF, (SIS-R) PF, GCSPF and HGCPF filters for nonlinear inference problems with non-Gaussian process noise under sparse measurements. Nevertheless a proper selection and optimization of the *filter model* is required for its better performances.

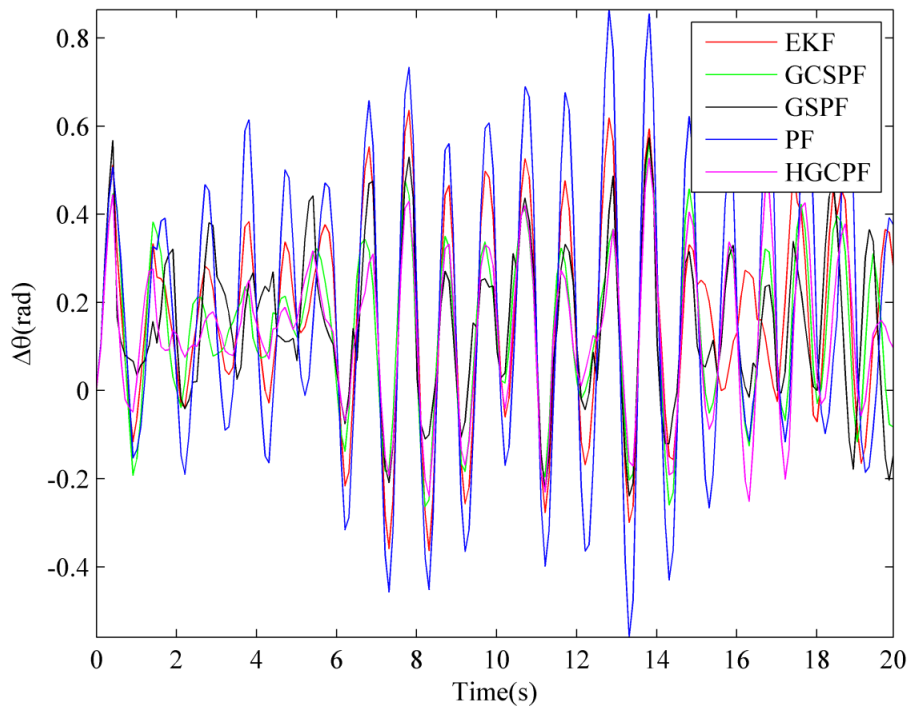


Figure 5-14: Comparison of time history of errors in angular position $\Delta\theta$ between true trajectory and estimated trajectory (wind gust simple pendulum model) for different filters.

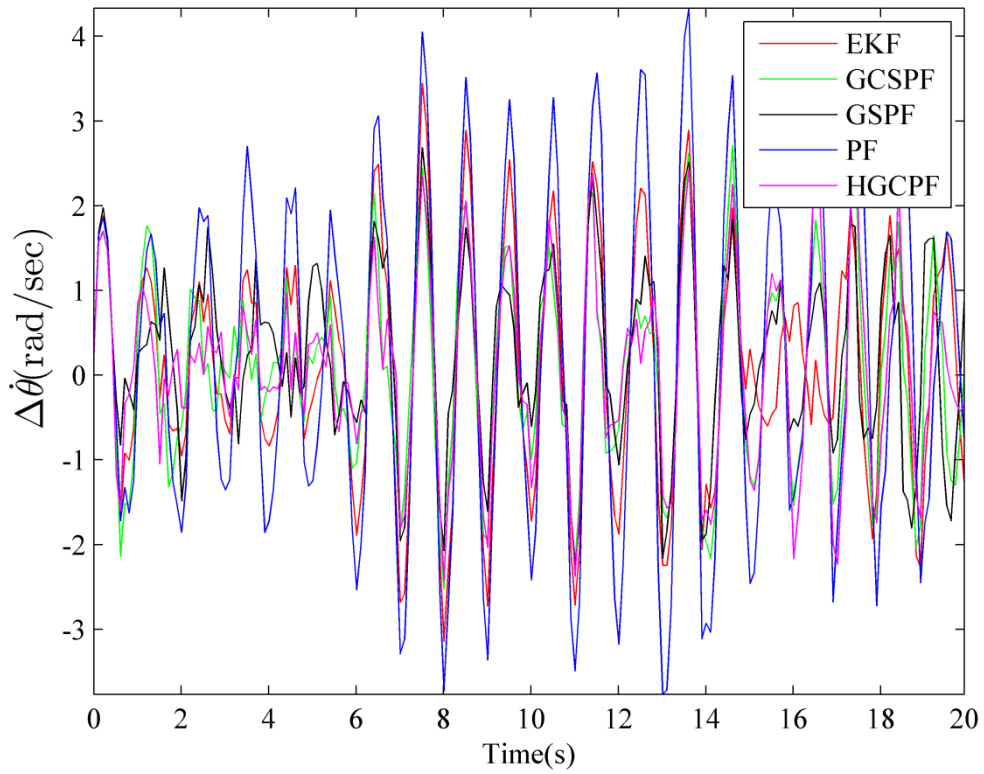


Figure 5-15: Comparison of time history of errors angular velocity $\Delta\dot{\theta}$ between true trajectory and estimated trajectory (wind gust simple pendulum model) for different filters.

Filter	Mean Error (ME)		Root Mean Square Error (RMSE)	
	$\Delta\theta(\text{rad})$	$\Delta\dot{\theta}(\text{rad. sec}^{-1})$	$\Delta\theta(\text{rad})$	$\Delta\dot{\theta}(\text{rad. sec}^{-1})$
EKF	-0.1655	-0.1138	0.2773	1.3500
GCSPF	-0.1382	-0.0909	0.2273	1.1412
HGCPF	-0.1329	-0.1330	0.2149	1.0884
GSPF	-0.1529	-0.2365	0.2342	1.0606
PF	-0.1736	-0.1714	0.3687	1.9866

Table 5-12: Comparison of filtering performances in terms of ME and RMSE for wind gust simple pendulum model.

5.7.3 Experiment – Radar Based Orbit Determination

In this experiment an orbit determination of a satellite through radar measurements is being looked into. It has been described in Chapter: 3 and 4 that satellite dynamics around our planet are highly nonlinear functions of its state (position and velocity) variables. The equations of motion of a satellite around a non-spherical geopotential in ECI reference frame [13] are rewritten in the form:

$$\dot{\mathbf{r}} = \mathbf{v} \quad (5.83)$$

$$\dot{\mathbf{v}} = -\frac{\mu_E}{r^3} \mathbf{r} + \mathbf{a}_G \quad (5.84)$$

where, $\mathbf{r} = (X, Y, Z)^T$, is the position vector, $\mathbf{v} = (\dot{X}, \dot{Y}, \dot{Z})^T$, is the velocity vector, both in ECI coordinates, and \mathbf{a}_G express here as higher zonal gravitational perturbation terms obtained by taking the gradient of potential function \mathcal{U} given in Equation: 3.30.

Equations: 5.83 and 5.84 with perturbation accelerations \mathbf{a}_G for Earth zonal harmonics up to J_4 are considered as the *true model* (see Equation: 4.40 to 4.42 for accelerations expressions in ECI reference frame). Given some specific initial conditions $\mathbf{x}_0 = [r_0 \ v_0]^T$ these equations are integrated using ODE Runge-Kutta 4 (RK-4) to get time history of position and velocity.

Now the measurement system is described of a satellite. The ECI position vector of satellite is related with the radar range vector and radar site vector through the following equation [13]:

$$\mathbf{r} = \mathbf{R}_s + \boldsymbol{\rho} \quad (5.85)$$

where, \mathbf{r} is the ECI coordinates of satellite, \mathbf{R}_s is the ECI coordinates of radar site, and $\boldsymbol{\rho}$ is the range vector from radar site to satellite.

The range vector $\boldsymbol{\rho}$ from the radar site to the satellite is described in Topocentric coordinate system (see Figure: 5-16 for illustration) in terms of the “zenith”, “east” and “north” as:

$$\boldsymbol{\rho} = \rho_u \hat{\mathbf{u}} + \rho_e \hat{\mathbf{e}} + \rho_n \hat{\mathbf{n}} \quad (5.86)$$

The range can be obtained as:

$$\rho = \sqrt{\rho_u^2 + \rho_e^2 + \rho_n^2} \quad (5.87)$$

The azimuth (az) and elevation (el) angles are expressed by:

$$az = \tan^{-1} \left(\frac{\rho_e}{\rho_n} \right) \quad (5.88)$$

$$el = \tan^{-1} \left(\frac{\rho_u}{\sqrt{\rho_e^2 + \rho_n^2}} \right)$$

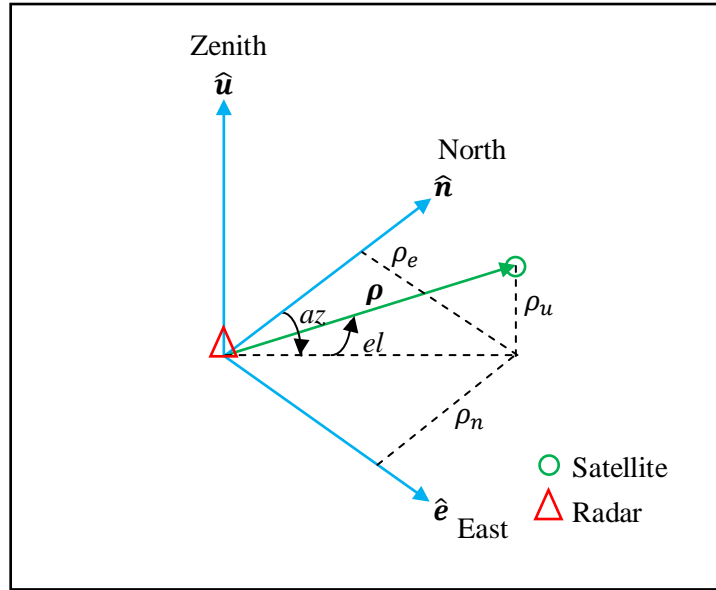


Figure 5-16: Measurement model description in Topocentric Coordinate System.

The east, north, and zenith unit vectors in Topocentric coordinate system is given by [13]:

$$\hat{e} = \begin{pmatrix} -\sin \Lambda \\ \cos \Lambda \\ 0 \end{pmatrix}, \hat{n} = \begin{pmatrix} -\sin \phi \cos \Lambda \\ -\sin \phi \sin \Lambda \\ \cos \phi \end{pmatrix}, \hat{u} = \begin{pmatrix} \cos \phi \cos \Lambda \\ \cos \phi \sin \Lambda \\ \sin \phi \end{pmatrix} \quad (5.89)$$

where, ϕ and Λ are geographical latitude and longitude of radar site respectively. By defining the orthogonal transformation as:

$$\mathfrak{N} = (\hat{\mathbf{e}} \hat{\mathbf{n}} \hat{\mathbf{u}})^T \quad (5.90)$$

The satellite's Topocentric coordinates in terms of radar site latitude and longitude may be obtained through following transformation:

$$\boldsymbol{\rho} = \begin{pmatrix} \rho_e \\ \rho_n \\ \rho_u \end{pmatrix} = \mathfrak{N}(\mathcal{R}_z(\Theta))\mathbf{r} - \mathbf{R}_s \quad (5.91)$$

where, $\mathcal{R}_z(\cdot)$ stands for the rotation about z-axis and Θ = Greenwich Mean Sidereal Time (GMST) [13]. GMST is also termed as Greenwich hour angle which denotes the angle between the mean vernal equinox (see Figure: 3-3) of date and the Greenwich meridian (see Figure: 3-5). It is a direct measure of Earth's rotation and expressed in angular units as well as time. For example 360 degrees (2π) correspond to 24 hours. Time calculations for satellite orbit predictions and determination are usually carried out in Julian Date (JD) [13],[12] due to its continuous nature. A Julian Date (JD) is the number of days since noon 1 January, 4713 BC including the fraction of day. Presently, the JD numbers are already quite large therefore a Modified Julian Date is defined as: $\text{MJD} = \text{JD} - 2400000.5$ (see ref [13] for computation of GMST from MJD / UTC). The *filter model* is defined for acquiring the *estimated trajectory* on the same lines as discussed for the simple pendulum model [116]:

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}, t)dt + \mathbf{G}(\mathbf{x}, t)d\boldsymbol{\beta}(t) \quad (5.92)$$

$$\mathbf{x} = \begin{bmatrix} X \\ Y \\ Z \\ \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix}, \mathbf{f}(\mathbf{x}, t) = \begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \\ -\frac{\mu_E X}{r^3} + J_2(X) \\ -\frac{\mu_E Y}{r^3} + J_2(Y) \\ -\frac{\mu_E Z}{r^3} + J_2(Z) \end{bmatrix}$$

$$\mathbf{G}(\mathbf{x}, t) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \sigma_{41} & 0 & 0 \\ 0 & \sigma_{52} & 0 \\ 0 & 0 & \sigma_{63} \end{bmatrix}, d\boldsymbol{\beta}(t) = \begin{bmatrix} d\beta_1(t) \\ d\beta_2(t) \\ d\beta_3(t) \end{bmatrix}$$

where, $J_2(X)$, $J_2(Y)$ and $J_2(Z)$ is the zonal perturbation term until J_2 for ECI X, Y and Z coordinates (see Equation: 4.40 for their mathematical expressions), $\mathbf{G}(\mathbf{x}, t)$ is the dispersion matrix $d\boldsymbol{\beta}(t)$ is the Brownian motion increment vector, σ_{41}, σ_{52} and σ_{63} are diffusion coefficients and subscripts denote row and column respectively, and $d\beta_1, d\beta_2$ and $d\beta_3$ are individual scalar Brownian motion increment (explained in Section: 5.7).

The Eglin US Air Force Base (AFB) is selected as radar site with $\phi = 30.2316$ deg and $\Lambda = 86.2147$ deg W. Each measurement consists of range, azimuth and elevation angles and the measurement errors were considered to be Gaussian distributed with following variances (adapted from ref [59]):

$$\sigma_{\text{range}} = 25 \text{ m}, \sigma_{\text{azimuth}} = 0.015 \text{ deg}, \sigma_{\text{elevation}} = 0.015 \text{ deg} \quad (5.93)$$

Initial conditions for satellite to generate *true trajectory* are [59] :

$$a = 6981.018 \text{ km} \quad (5.94)$$

$$e = 7.535 \times 10^{-4}$$

$$I = 51.60 \text{ deg}$$

$$\Omega = 25.003 \text{ deg}$$

$$\omega = 339.31 \text{ deg}$$

$$M = 57.04 \text{ deg}$$

IC to filter with large position variances as 10^5 m^2 and velocity variances $500 \text{ m}^2 \cdot \text{sec}^{-2}$ are [59]:

$$\hat{\mathbf{P}}_0 = \text{diag}([10^5 \ 10^5 \ 10^5 \ 500 \ 500 \ 500]) \quad (5.95)$$

$$\hat{\mathbf{x}}_0 = [\mathbf{x}_0] + \left[\sqrt{10^5} \ \sqrt{10^5} \ \sqrt{10^5} \ \sqrt{500} \ \sqrt{500} \ \sqrt{500} \right]^T$$

In general the IC for mean and higher order moments is obtained using algorithms for initial OD (IOD) of satellite such as nonlinear least squares or Herricks-Gibbs (HG) methods [12]. For simulation purposes the IC in Equation: 5.95 are acquired from reference [59] which is based on HG method. Keeping in line with the experimental setup of the simple pendulum under sparse measurements, availability of measurements for $\approx 3\%$ of the orbital period of satellite ($T = 96.75 \text{ min}$) is selected. Therefore, the observations are recorded for 3 minutes per orbital period with a 5 sec gap between the measurements. The filtering assessment criteria are kept the same (Equations:

5.78 to 5.80). The time history of IE in ECI coordinates are given in Figures: 5-17 to 5-19. The number of particles for each filter is selected as 250 to produce these figures. However, experiments with different numbers of particles were carried out and the results indicated almost the same comparison ratios between the filters.

Filter	RMSE (Position) meters			RMSE(Velocity) meters/s		
	X	Y	Z	\dot{X}	\dot{Y}	\dot{Z}
GCSMPF	81.06	112.43	116.89	3.86	4.35	5.15
GSPF	97.05	136.77	191.01	5.02	4.53	6.16
PF	589.22	539.80	304.94	16.17	11.50	11.44

Table 5-13: Comparison of filters Root Mean Square Errors (RMSE)

Filter	ME (Position)[m]			ME(Velocity)[m/s]		
	X	Y	Z	\dot{X}	\dot{Y}	\dot{Z}
GCSMPF	-20.59	88.27	-105.59	-0.79	-0.07	-0.94
GSPF	48.78	70.93	-147.83	-1.65	0.93	-2.09
PF	-435.99	473.12	-242.65	-4.38	-0.91	-3.53

Table 5-14: Comparison of filters Mean Errors (ME)

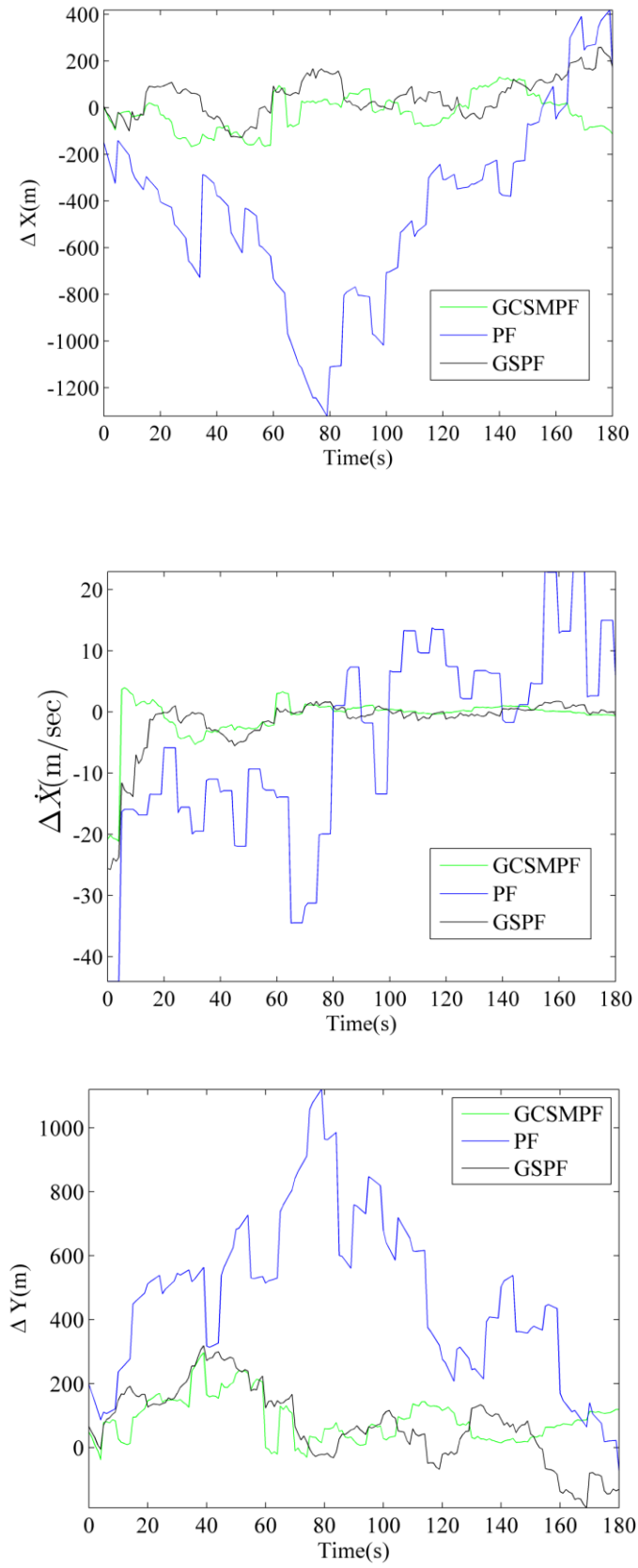


Figure 5-17: Time history of errors (Δ) in ECI X (top), \dot{X} (middle), and Y (bottom). The measurement frequency is 0.2 Hz.

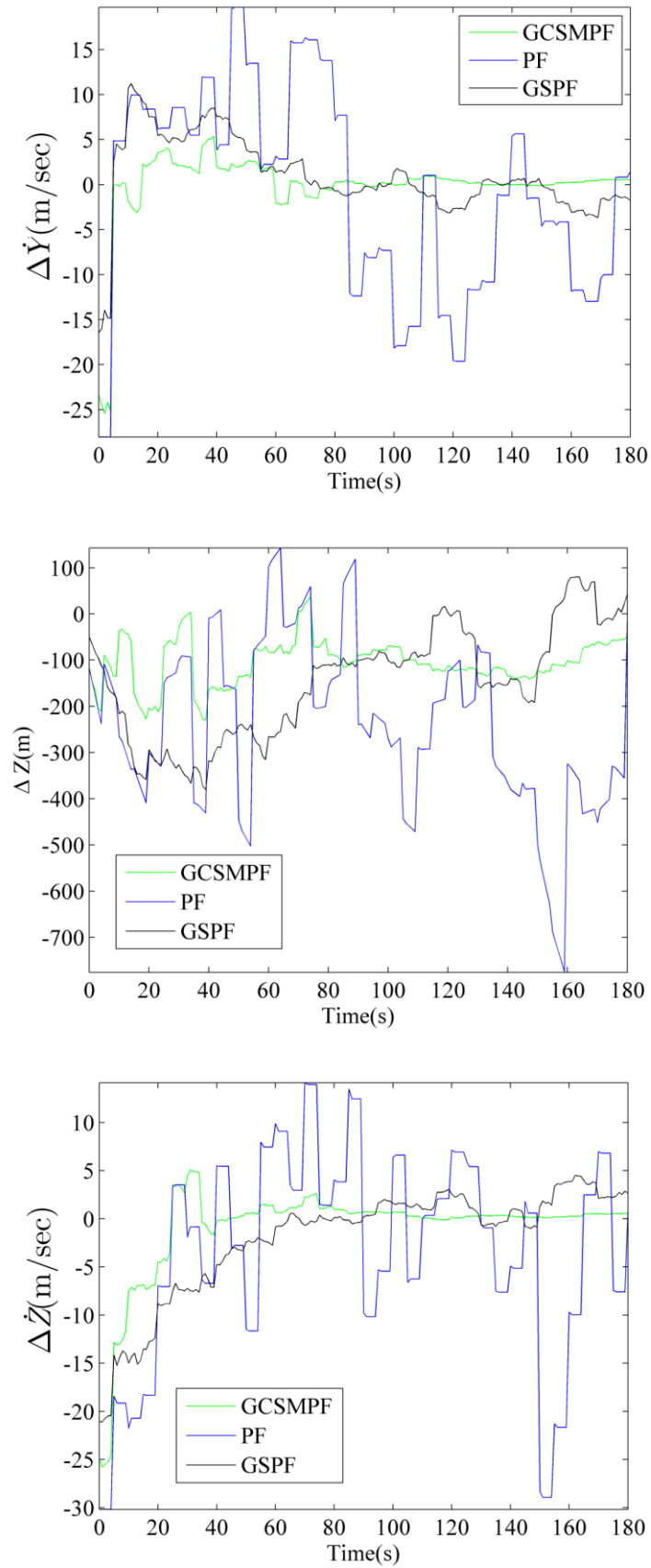


Figure 5-18: Time history of errors (Δ) in ECI \dot{Y} (top), Z (middle), and \dot{Z} (bottom). The measurement frequency is 0.2 Hz.

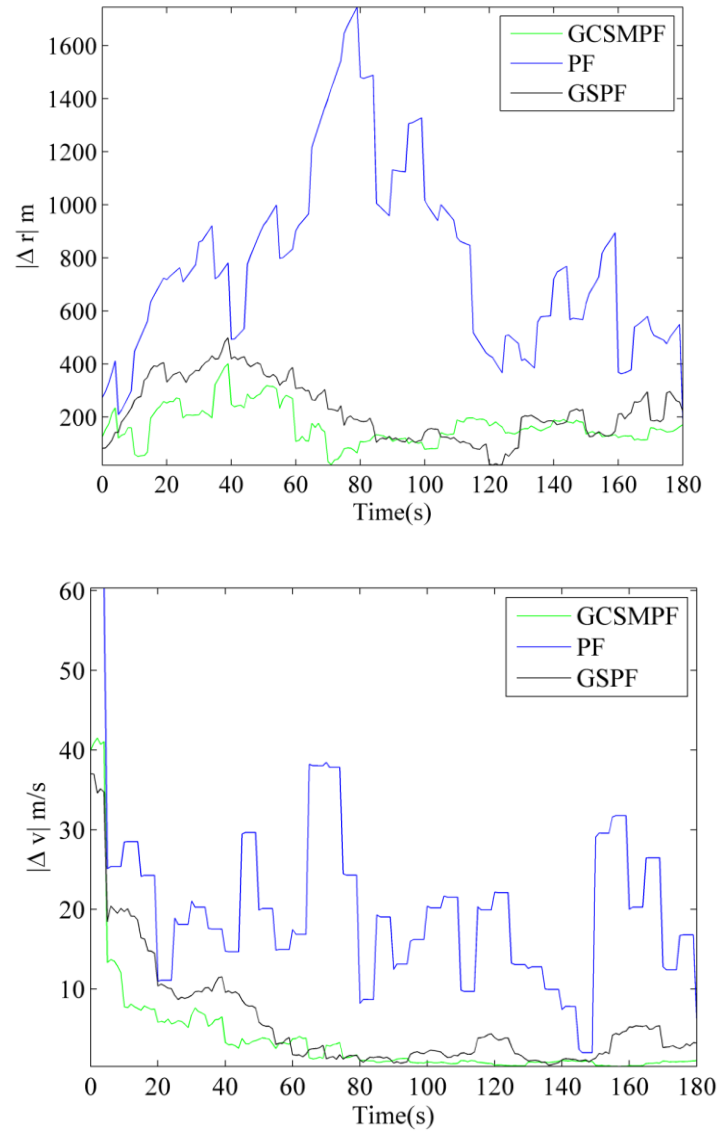


Figure 5-19: Time history of magnitude of errors in position $|\Delta r|$ m (top) and velocity $|\Delta v|$ m/s (bottom). Measurement frequency is 0.2 Hz.

The RMSE results for GCSMPF (Table: 5-13) are better than the other two types of filtering techniques. Copula type of random number generator (`coprangcs`) has been used in GCSMPF with two mixands (components). The GSPF has also been implemented with two Gaussian mixands. The increase in number of components for GCSMPF and GSPF in this experiment decreases computational speed. Moreover, it is further affected with increase in number of particles used by the algorithms. Therefore, two component mixtures PDF for GCSMPF and GSPF are considered for satellite OD experiment. This selection of number of mixture components provides comparable computational speed with respect to (SIS-R) PF. The results of ME errors (Table: 5-14) are away by

meters from zero mean consistency criteria for all the filters. Nevertheless, for this type of experimental setting these errors are still not very large. The plots in Figures: 5-17 to 5-19 are averaged results over several runs for each filter. Similar efforts were performed on each filter to perform optimally. Therefore, these results give an indication of different inference algorithms tested on a nonlinear problem. The deteriorated performance of the (SIS-R) PF is due to the large uncertainty provided to the filter (Equation: 5.95). The performance of the PF improves considerably if the values provided in Equation: 5.95 are reduced. However, reduced uncertainty may not be very realistic for space object estimation under sparse measurements scenario. After the termination of set of measurements the performance of filters will now be observed for second pass over the same radar site after one orbital period later. The duration of measurements is kept same i.e. three minutes. The estimates and associated uncertainties for each filter are computed by using propagated particles until the first observation of second orbital period. The time history of errors in ECI coordinates are shown in Figure: 5-20 to 5-22. The performance of the (SIS-R) PF is significantly suboptimal compared with the GSPF and GCSMPF. Therefore these figures only illustrate the later two filtering comparisons. Table 5-15 and 5-16 shows RMSE and ME, respectively.

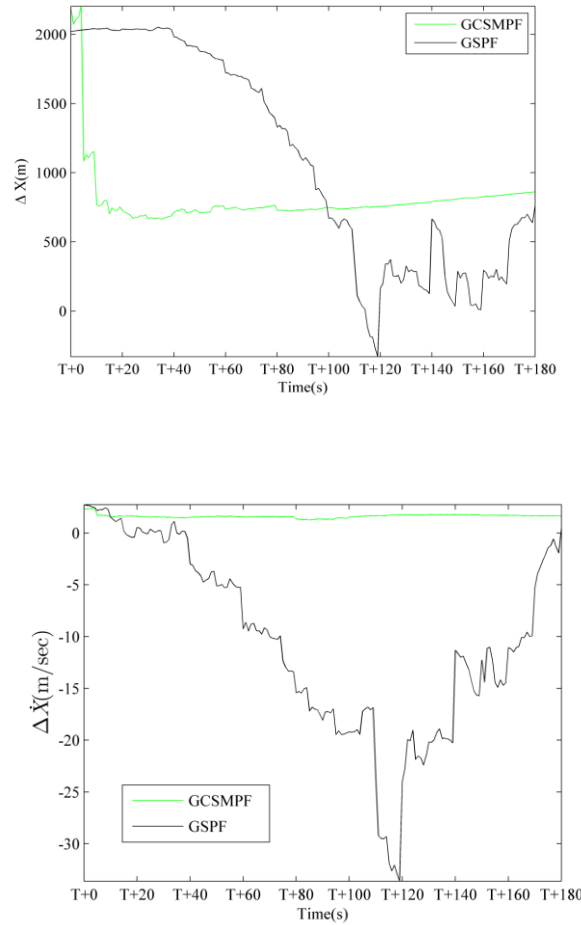


Figure 5-20: Time history of errors (Δ) in ECI X (m) and \dot{X} (m/s) after one orbital period T , where ($T = \sim 97$ min).

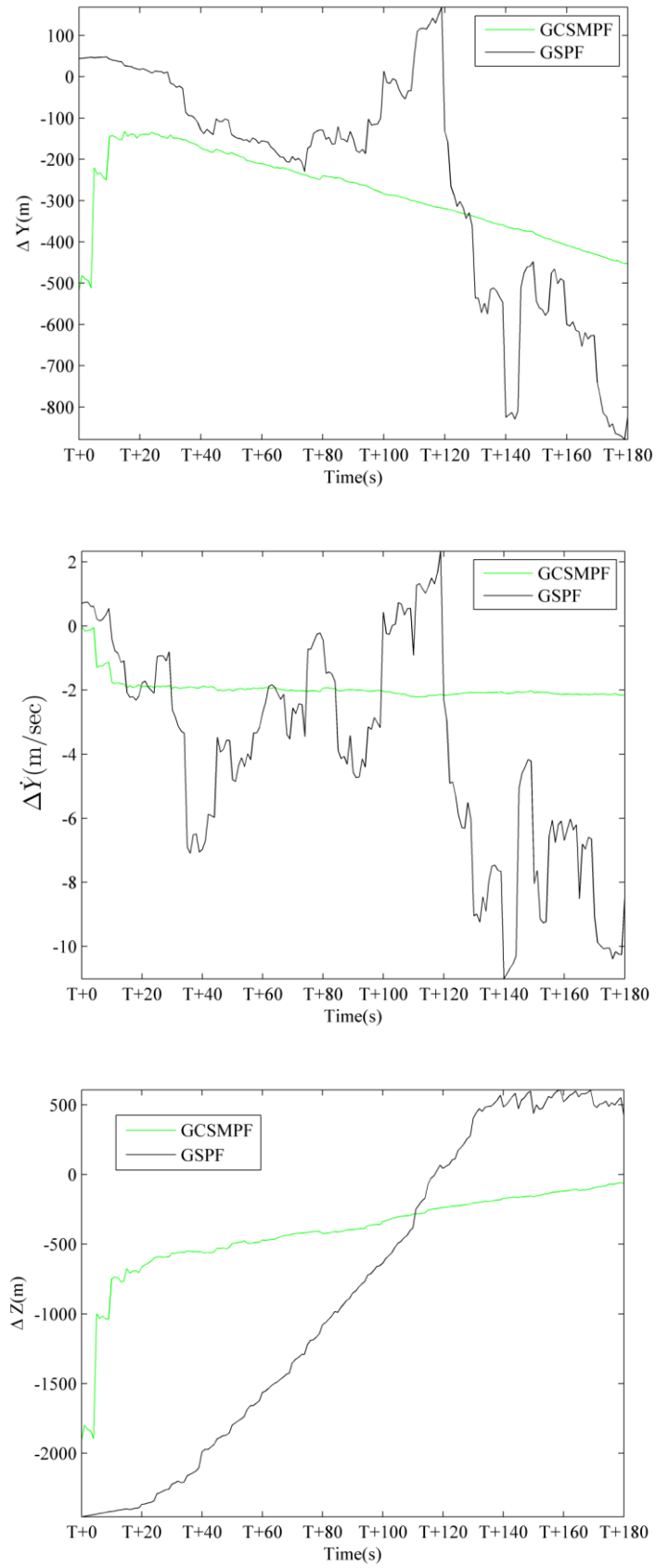


Figure 5-21: Time history of errors (Δ) in ECI Y (top), \dot{Y} (middle), and Z (bottom). The measurement frequency is 0.2 Hz after one orbital period T , where ($T = \sim 97$ min).

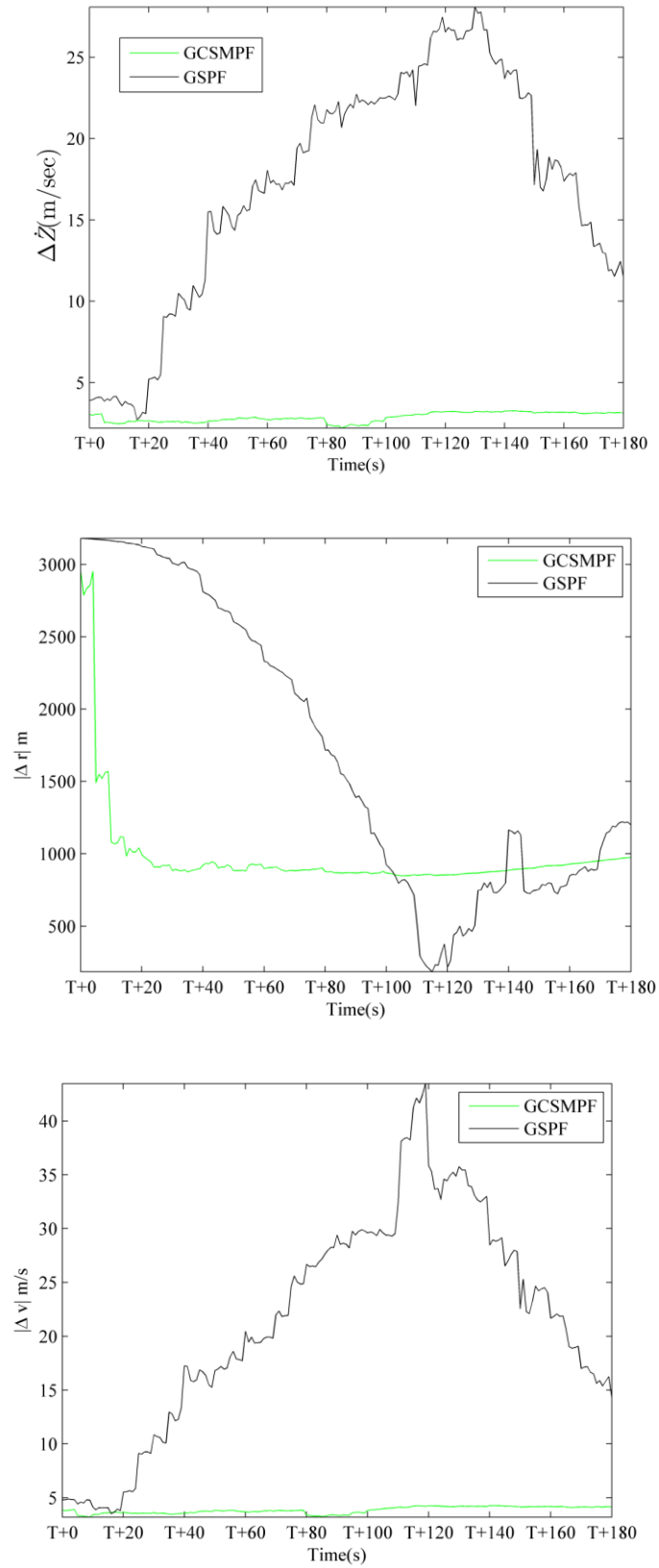


Figure 5-22: Time history of errors (Δ) in ECI \dot{Z} (top), $|\Delta r|$ (middle), and $|\Delta v|$ (bottom). The measurement frequency is 0.2 Hz after one orbital period T , where ($T = \sim 97$ min).

Filter	RMSE (Position)			RMSE(Velocity)		
	meters			meters/s		
	X	Y	Z	\dot{X}	\dot{Y}	\dot{Z}
GCSMPF	840	299	525	1.6	2.0	2.9
GSPF	1346	365	1401	13.8	5.2	18.6

Table 5-15: Comparison of filters Root Mean Square Errors (RMSE) after one orbital period

Filter	ME (Position)[m]			ME(Velocity)[m/s]		
	X	Y	Z	\dot{X}	\dot{Y}	\dot{Z}
GCSMPF	805	-282	-414	1.6	-2.0	2.9
GSPF	1099	-238	-842	-10.5	-3.9	17.2

Table 5-16: Comparison of filters ME after one orbital period

The results of sequential filters show improved efficiency of GCSMPF over GSPF, especially due to the lower RMSE / ME and better convergence. Note the significant divergence of the GSPF, which can lead to increased errors in later orbital periods. In general, divergence of filters happens due to its very low error covariance matrix output (extremely good confidence in estimates), which can be erroneous. Therefore, one may initialize the covariance matrix as soon as the error covariance matrix goes less than a certain threshold. In order to illustrate this situation consider an example of ECI positional errors of GSPF and GCSMPF along with their covariance for a single set of observations over the same radar site. Figures: 5-23 to 5-26 shows time history of errors and positional covariance, respectively for GSPF and GCSMPF.

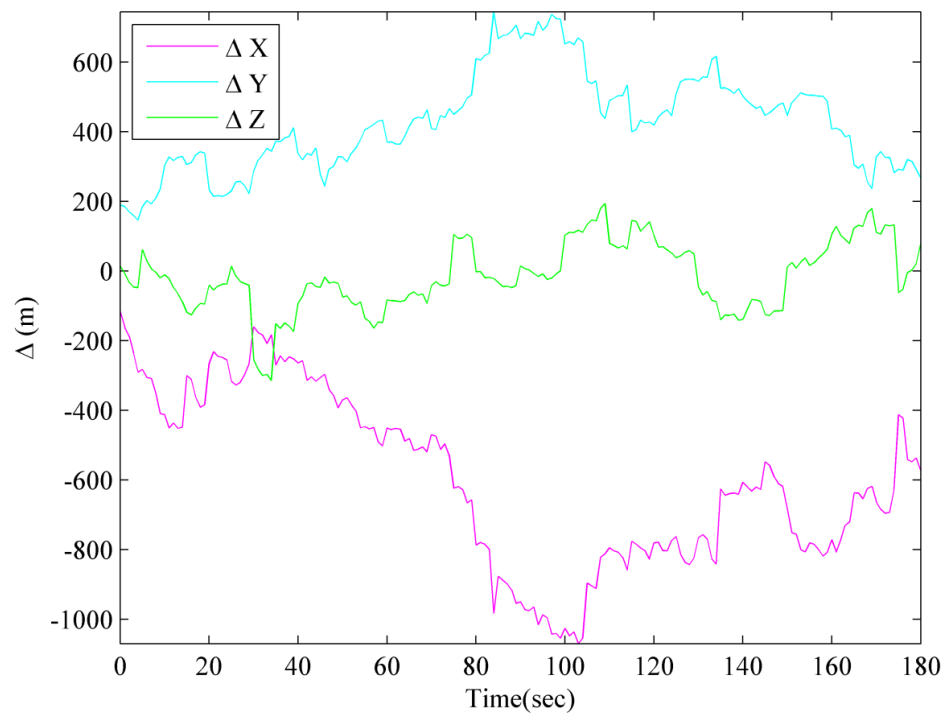


Figure 5-23: Time history of position errors (Δ) in ECI coordinates for a GSPF.

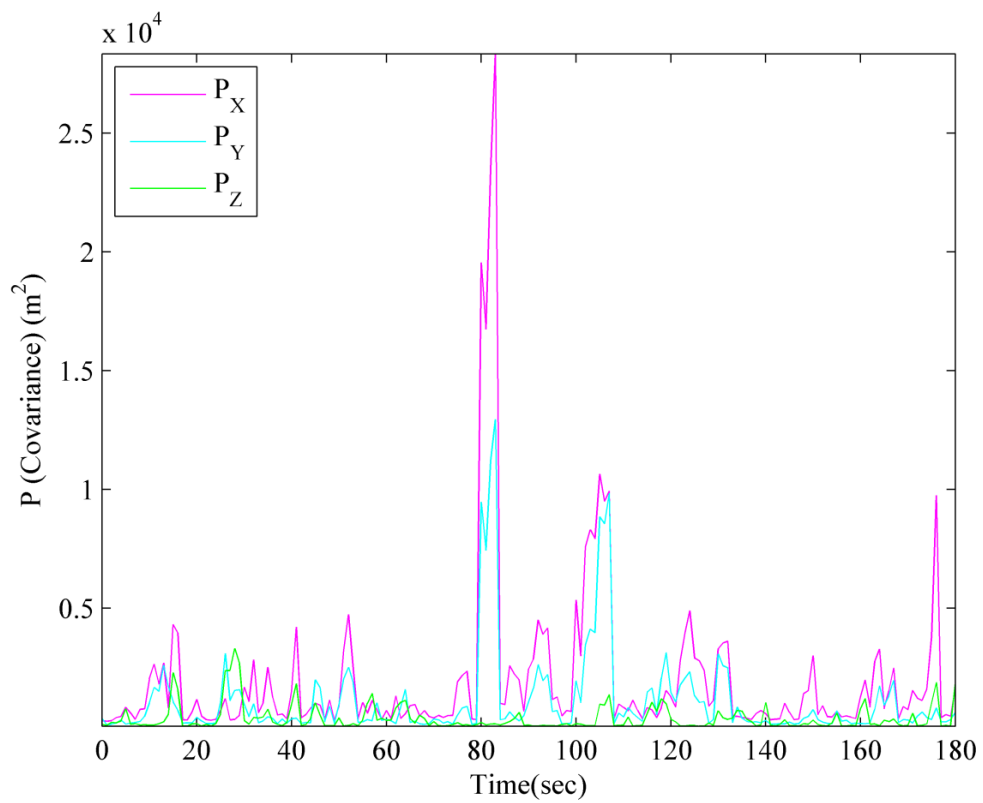


Figure 5-24: Time history of positional covariance for a GSPF. Note reduced (very low) covariance which may cause filter divergence as it assumes more confidence in estimates.

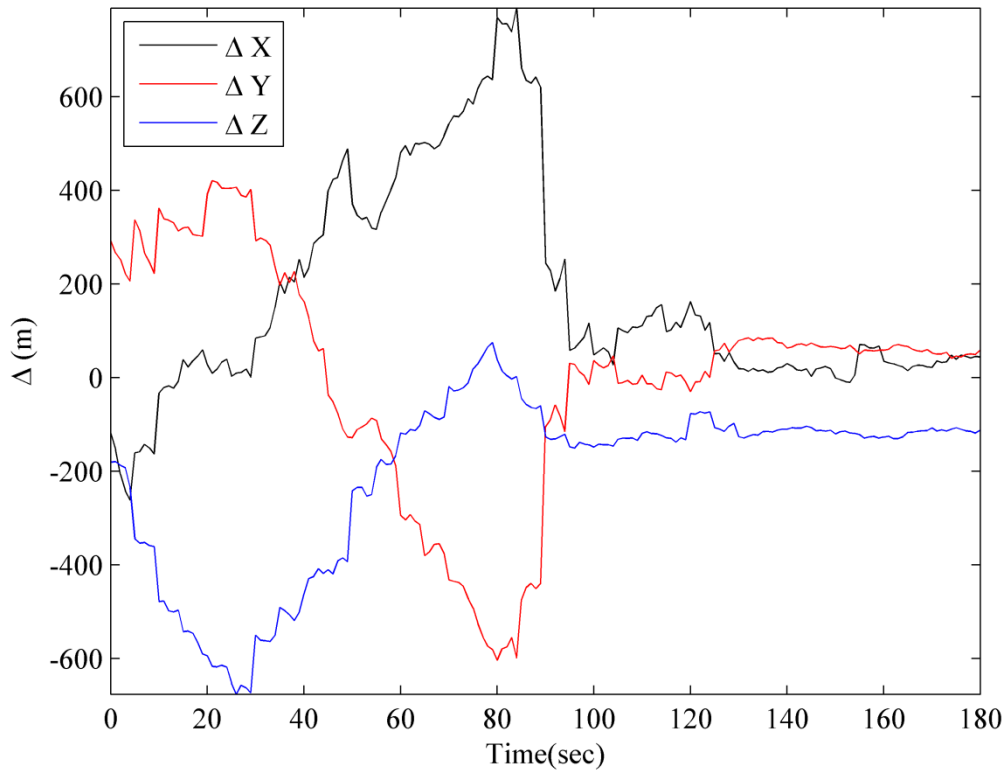


Figure 5-25: Time history of position errors (Δ) in ECI coordinates for a GCSMPF.

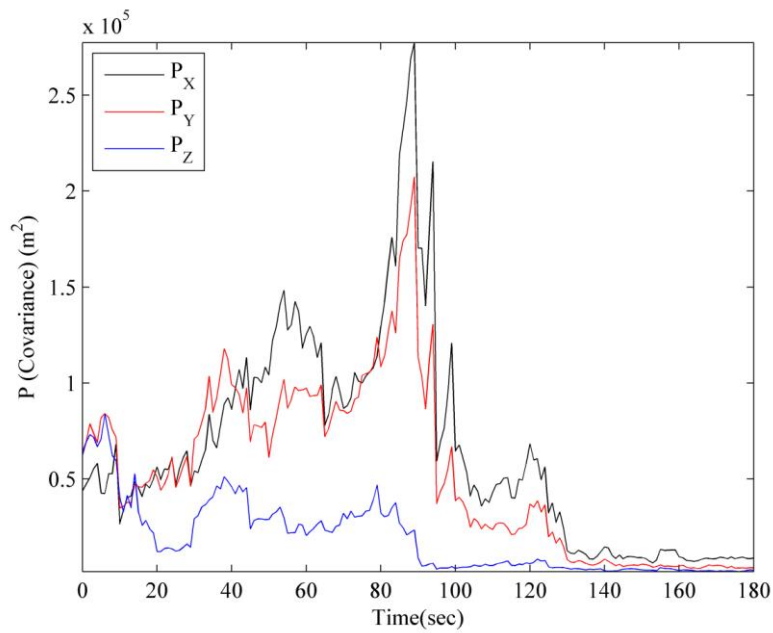


Figure 5-26: Time history of positional covariance for a GCSMPF.

One may clearly observe that the output covariance of GSPF is low compared to GCSMPF which means more confidence in estimates which could be erroneous. Hence, the GSPF is more likely to produce diverged estimates. In these experiments we have not considered covariance re-initializing in

order to evaluate filtering performances without using such engineering practice in order to gauge error trends. We now observe the performance of the GCSMPF for orbital periods greater than 2. Figures: 5-27 to 5-29 provides the time history of position errors until 7 orbital periods.

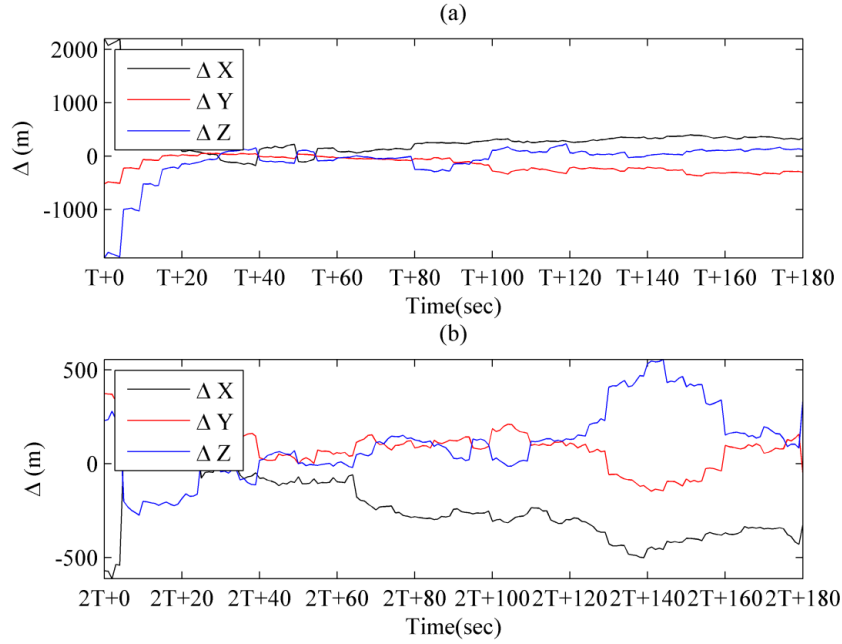


Figure 5-27: Time history of ECI position errors (Δ) for GCSMPF during subsequent orbital periods, (a) 2nd orbital period, (b) 3rd orbital period, where ($T = 5805$ sec).

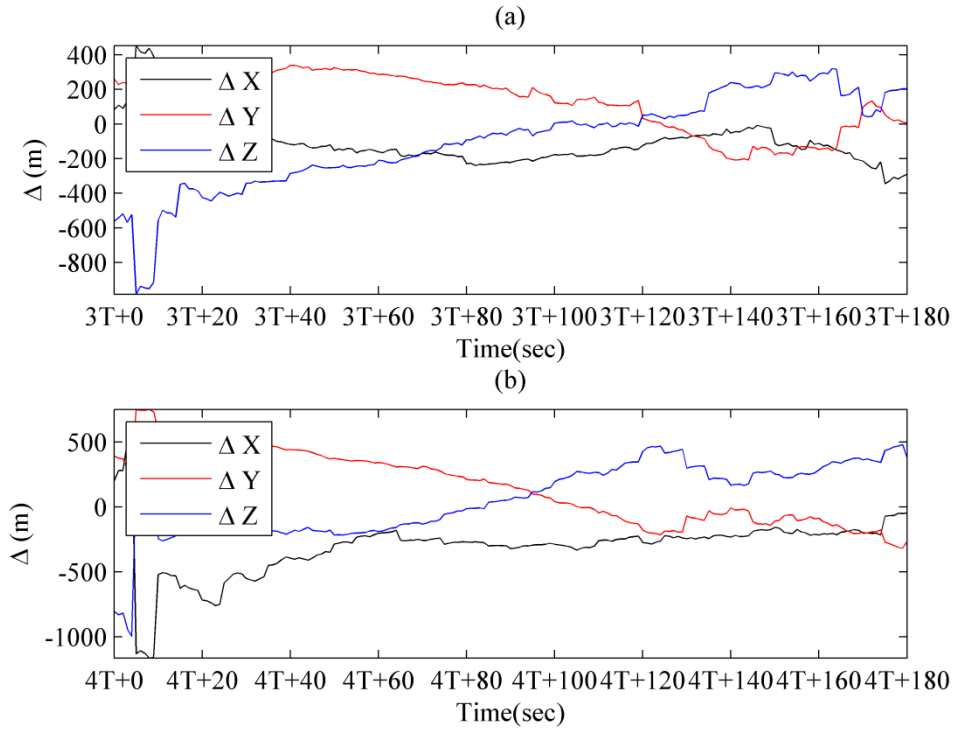


Figure 5-28: Time history of ECI position errors (Δ) for GCSMPF during subsequent orbital periods, (a) 4th orbital period, (b) 5th orbital period, where ($T = 5805$ sec).

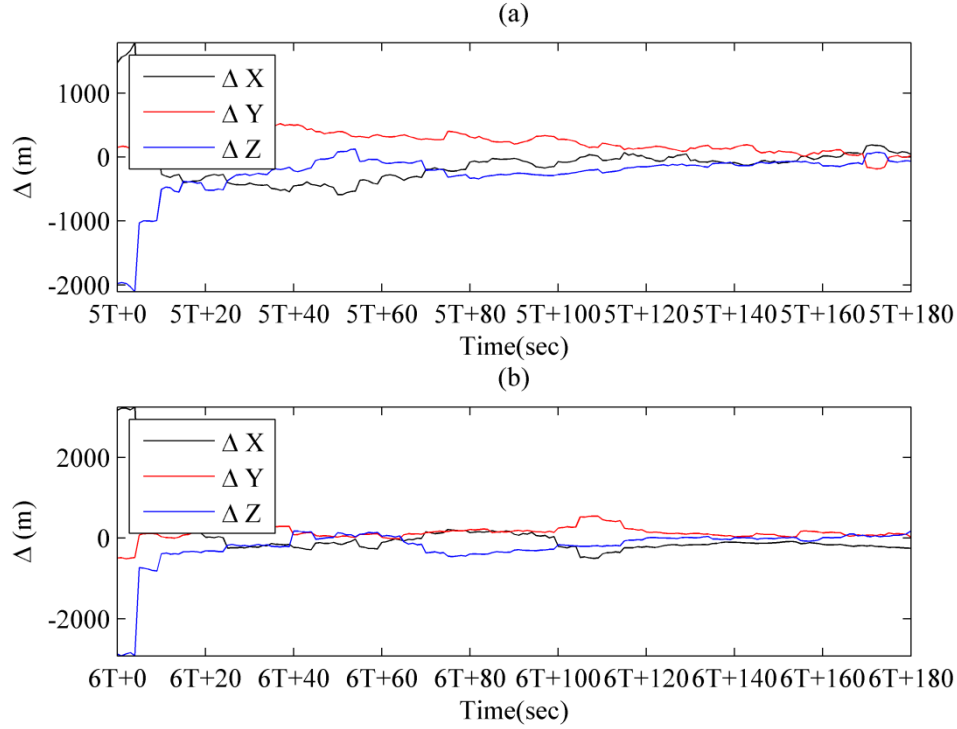


Figure 5-29: Time history of ECI position errors (Δ) for GCSMPF during subsequent orbital periods, (a) 6th orbital period, (b) 7th orbital period, where ($T = 5805$ sec).

The figures gives an indication of convergence and almost similar error statistics for the GCSMPF. Hence, the performance of GCSMPF over subsequent orbital periods is gives us indication of its suitability for use in satellite OD.

5.8 Summary

In this chapter a detailed description of PF and its variants have been discussed. The background necessary for understanding of PF methodology is elaborated in detail. New filters based on GCS namely GCSMPF, GCSMPF, and HGCPF have been presented. The algorithms have been compared with PF, GPF, GSPF and EKF for nonlinear simple pendulum and orbit determination through radar measurements. The results show improvements in IE, RMSE and ME for the new filters (GCSMPF, GCSMPF and HGCPF). GCS and its mixtures can be considered as better choice for replacement of Gaussian PDF in nonlinear filtering applications especially for improvement in particle filtering. An important aspect of filters based on higher order GCS and its mixture is computational complexity associated with generation of random numbers. In this chapter, AR and Gaussian copula based methods are used which may not be always optimal. For example, in AR method for bivariate GCS random vectors the rejected variates are approximately 20-30% (see Table: 5-2) which severely impacts speed of execution. Therefore, there is a need for development of better random number generator for GCS. In order to implement discrete-time filtering the continuous-time nonlinear

dynamical systems used in the experiments are discretized using a fixed time step of numerical integration method (RK-4). In general high fidelity numerical solution is obtained by keeping a very short time step (order of millisecond). This significantly affects the speed of execution in real time particle filtering for satellite OD which owes to high dimensionality and more number of particles used for such problems. However, GCSPF and GCSMPF can be implemented in parallel which makes it suitable for high speed Very Large Scale Integrated Circuit (VLSI) based implementation for real time filtering.

6 Development of Mixture Culver Filter

6.1 Introduction

In continuation to the work in Chapter: 5, which pertains to *discrete filtering*; this chapter describes the sequential state estimation known as *continuous-discrete filtering*. In *continuous-discrete filtering*, the evolution of a dynamical system is considered as a continuous process; whereas, the measurements are taken at discrete instants of time [6]. See Figure: 6-1 for description of a continuous process being observed at discrete time [21]. It is like a time series being observed at discrete instants.

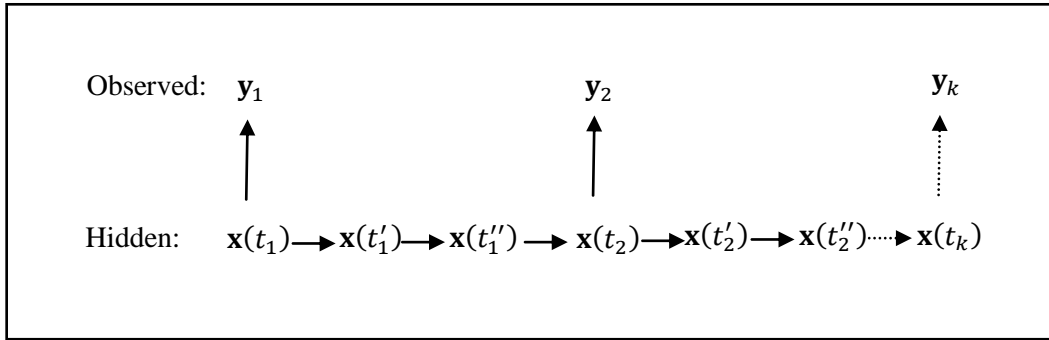


Figure 6-1: In *continuous-discrete filtering* evolution of time is continuous and measurements are taken at discrete instants. The progression of time from one measurement until another measurement is continuous e.g., $(t_1 \rightarrow t'_1 \rightarrow t''_1 \rightarrow t_2)$.

As seen in Chapter: 5, filtering of continuous time nonlinear dynamical system (Equation: 5.69, 5.81, 5.83 and 5.84) would require formulation of appropriate mathematical model for evolution of the state of the system along with its uncertainties i.e., the state predictive PDF. Since the evolution in time is a continuous process therefore dynamical systems can be more realistically represented as SDE (first of Equation: 5.73). The advantage of *continuous-discrete filtering* is that the sampling interval ($\Delta t = t_k - t_{k-1}$) can change between the measurements unlike discrete filtering where sampling time should be constant [21]. Nevertheless, the mathematical model for measurement system is identical to

the *discrete filtering* (Equation: 5.2). The sequential state estimation of such a system can also be realized using Bayes' formula as given in Equation: 5.4. However, the prediction of state transition PDF for continuous time dynamical system satisfies the FPKE. It is a linear parabolic type PDE expressed as [1],[5],[6]:

$$\frac{\partial p}{\partial t} = - \sum_{i=1}^d \frac{\partial}{\partial x_i} (\mathbf{f}(\mathbf{x}(t), t)p) + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \frac{\partial^2}{\partial x_i \partial x_j} [(\mathbf{G}\mathbf{Q}\mathbf{G}^T)_{ij}p] \quad (6.1)$$

where, p is the state PDF, $\mathbf{f}(\mathbf{x}(t), t)$ is the nonlinear function modelled in continuous time also known as drift or advective term, \mathbf{G} is the dispersion matrix, $\mathbf{G}\mathbf{Q}\mathbf{G}^T$ is the diffusion matrix of SDE, \mathbf{Q} is the diffusion matrix for white Gaussian noise and d is the dimension of the system.

The PDE in Equation: 6.1 for nonlinear dynamic systems can be solved using numerical integration methods such as Finite differencing (FD) [78] wherein, the Bayes' *a posteriori* PDF at time step t_{k-1} is propagated forward to obtain state transition PDF of the system until the measurement is received at time step t_k . On receiving the measurement, the Bayesian *update* step is realized using the Bayes' formula given in Equation: 5.4. The block description of *continuous-discrete filtering* is shown in Figure: 6-2.

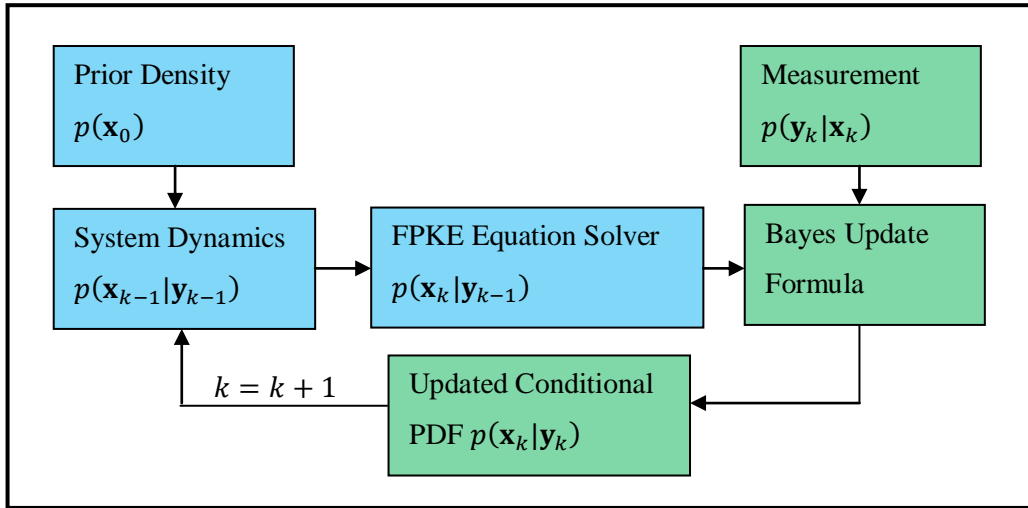


Figure 6-2: The block description of *continuous-discrete filtering*. Any optimal estimates of state such as MMSE can be obtained from updated conditional PDF $p(\mathbf{x}_k|\mathbf{y}_k)$.

The main complexity in obtaining the optimal solution (see Figure: 6-2) of nonlinear *continuous-discrete filtering* problem arises due to the need for solution of Equation: 6.1. An analytical solution of

Equation: 6.1, is usually possible for linear dynamical systems [5]. In general, numerical methods are required to solve it for nonlinear systems of lower dimensions (dimensions ≤ 6) due to recent increase in computational resources [25],[5]. However, sequential state estimation using numerical solution of PDE is not considered optimal [25]. This is due to the requirement of enormous computational resources in terms of storage of PDF which amounts to an infinite dimensional vector, and prohibitively large processing times. Therefore, numerical solution of FPKE for *continuous-discrete* state estimation of satellite orbital dynamics is not practicable especially, in satellite onboard Orbit Determination (OD) systems such as GPS [117], Inertial Navigation System (INS) [117] or celestial navigation systems [118],[1]. This is mainly due to less computational resources available in satellite On-Board-Computers (OBC) [14]. Furthermore, even for ground based sequential satellite OD systems using radar measurements, numerical solution of the FPKE poses excessive memory requirements, owing to storage and recursion of entire Bayes' *a posteriori* PDF at each time step. Hence, there is a need for computationally tractable solutions to Bayes' *a posteriori* PDF to enable practicable in orbit, and ground based, satellite OD systems.

The state transition PDF of a continuous time nonlinear dynamical system can be approximated by its first few moments, such as in EKF the state transition PDF is approximated as Gaussian; therefore, only its first two moments i.e., the mean and covariance are propagated forward between the measurements using linearized dynamics [5],[6]. In general, more accurate representation of state transition PDF can be obtained using its higher order moments, such as in third order Culver Filter (CF) [1], where moments up to third order are used in GCS approximation of state transition PDF.

The most commonly used sequential nonlinear filter for satellite OD, is the EKF which approximates the Bayes' *a posteriori* PDF as Gaussian [13],[59],[14],[117],[53]. However, as shown in Chapter: 1 and 5, the mixture formulation of the PDFs (Gaussian and GCS) i.e., GMM and GCSMM are better alternatives to approximate non-Gaussian PDFs. Therefore, in this chapter one would consider an extension of the EKF, based on the GMM approximation of the Bayes' *a posteriori* PDF for satellite OD. This nonlinear filter is commonly known as Gaussian Sum Filter (GSF) [44]. Together one would consider these two filters (EKF and GSF) as a *Kalman Filter* framework.

Another viable solution for nonlinear satellite OD in *continuous-discrete filtering* setup is CF. As briefly explained in Chapter: 1, 2 and 5, CF approximates Bayes' *a posteriori* PDF, as a third order GCS [1]. In CF, linearization of nonlinear dynamic and measurement function is done, respectively up to second and first order in Taylor series. Therefore, moments up to third order (i.e., mean, covariance and coskewness) are propagated forward (between the measurements) using linearized dynamics [1],[6]. However, the CF provides an exact optimal MMSE solution to the nonlinear Bayesian filtering problem under the assumption of third order GCS approximation of Bayes' *a posteriori* PDF and differentiability of nonlinear dynamics. As described in Chapter: 1 and 5, a lower order (order ≤ 4) GCS is suboptimal representation of true non-Gaussian PDF (see Figure: 5-4 and 5-5). Therefore

this dissertation proposes a new nonlinear filter based on GCSMM of lower order (order = 3) for an optimal approximation of the Bayes' *a posteriori* PDF for *continuous-discrete filtering* for nonlinear dynamical systems in general, and satellite OD in particular. This would serve as an enhancement of the original CF, and is called as Mixture Culver Filter (MCF). Together with CF and MCF one would form a *Culver Filter* framework. In the subsequent sections Culver Filter framework would be described in detail and applications pertaining to onboard navigation for lunar orbiter and OD of LEO satellite through ground based radar and GPS system would be evaluated under sparse measurements availability and highly uncertain initial conditions. The dynamic and measurement function for OD of LEO satellite based on radar measurements and lunar optical navigation are both highly nonlinear functions of their state variables. Therefore, a simple pendulum model of Chapter: 5 would not be utilized here as its measurement function is linear (see Equation: 5.71). A comparison of *Kalman and Culver Filter* framework would also be performed for the above mentioned experiments.

6.2 Continuous Discrete Nonlinear Filtering Problem

Consider a continuous time dynamical system expressed by the nonlinear Ito Stochastic Differential Equation (SDE) of the following form [1][6]:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t) + \mathbf{G}(\mathbf{x}(t), t)\mathbf{W}(t) \quad (6.2)$$

where, $\mathbf{x}(t) \in \mathbb{R}^d$ is the d -dimensional state of the stochastic process, $\mathbf{f}(\mathbf{x}(t), t) \in \mathbb{R}^{d \times 1}$ is the drift function of $\mathbf{x}(t)$ and t describing the system dynamics, $\mathbf{G}(\mathbf{x}(t), t) \in \mathbb{R}^{d \times m}$ is the dispersion matrix of function of $\mathbf{x}(t)$ and t , and $\mathbf{W}(t) \in \mathbb{R}^m$ is the white noise. However, considering white noise as derivative of Brownian motion $\left(\mathbf{W}(t) = \frac{d\boldsymbol{\beta}(t)}{dt}\right)$ [21][114], a more useful form of Equation: 6.2 can be written as [1]:

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t)dt + \mathbf{G}(\mathbf{x}(t), t)d\boldsymbol{\beta}(t) \quad (6.3)$$

where, $\boldsymbol{\beta}(t) \in \mathbb{R}^m$ is Brownian motion of mean equals to zero, and diffusion $\mathbb{Q}(t)$:

$$E[d\boldsymbol{\beta}(t)d\boldsymbol{\beta}(t)^T] = \mathbb{Q}(t)dt \quad (6.4)$$

Consider measurements are observed at discrete time t_k expressed as [8]:

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}(t_k)) + \mathbf{v}_k \quad (6.5)$$

where, $\mathbf{y}_k \in \mathbb{R}^q$ is the q -dimensional observation vector, $\mathbf{h}(\mathbf{x}(t_k)) \in \mathbb{R}^{q \times 1}$ = measurement function and $\mathbf{v}_k \in \mathbb{R}^q$ is the q -dimensional zero mean Gaussian noise process. The covariance of measurement noise is given by:

$$E[\mathbf{v}_k \mathbf{v}_j^T] = \delta_{ij} \mathbf{R}_k \quad (6.6)$$

where, δ_{kj} = Dirac delta function and \mathbf{R}_k = covariance matrix. It is assumed that the initial states of the dynamic system (Equation: 6.2) and measurement noise are independent. The problem is to find out state estimates conditioned on the measurements \mathbf{y}_k in the MMSE sense. Considering that the prior PDF of the dynamic system expressed in Equation: 6.2 is available and is continuously differentiable once with respect to t and twice differentiable with respect to $\mathbf{x}(t)$, then it can be shown that, between the observations, the conditional PDF $p(\mathbf{x}(t_k)|\mathbf{y}_{k-1})$ satisfies FPKE (Equation: 6.1). On receipt of measurement at t_k the conditional PDF known as Bayes' *a posteriori* PDF is computed using following expression [5],[33]:

$$p(\mathbf{x}(t_k)|\mathbf{y}_k) = \frac{p(\mathbf{y}_k|\mathbf{x}(t_k))p(\mathbf{x}(t_k)|\mathbf{y}_{k-1})}{\int_{-\infty}^{+\infty} p(\mathbf{y}_k|\mathbf{x}(t_k))p(\mathbf{x}(t_k)|\mathbf{y}_{k-1})d\mathbf{x}(t_k)} \quad (6.7)$$

where, $p(\mathbf{y}_k|\mathbf{x}(t_k))$ is given by:

$$p(\mathbf{y}_k|\mathbf{x}(t_k)) = \frac{1}{|2\pi\mathbf{R}_k|^{1/2}} \exp \left\{ -\frac{1}{2} [\mathbf{y}_k - \mathbf{h}(\mathbf{x}(t_k))]^T \mathbf{R}_k^{-1} [\mathbf{y}_k - \mathbf{h}(\mathbf{x}(t_k))] \right\} \quad (6.8)$$

Equation: 6.1 and 6.7 can be considered as predictor and corrector method for evolution of PDF. The mean of Bayes' *a posteriori* PDF (Equation: 6.7) gives the optimal state of the system in MMSE sense [5],[33] given by:

$$\hat{\mathbf{x}}_k = \int_{-\infty}^{+\infty} \mathbf{x}(t_k) p(\mathbf{x}(t_k) | \mathbf{y}_k) d\mathbf{x}(t_k) \quad (6.9)$$

6.3 Culver Filter

Restricted Gaussian assumption of state transition (Equation: 6.1) and Bayes' *a posteriori* PDFs (Equation: 6.7) by EKF may provide suboptimal state estimates [5],[48]. The first two moments provided by EKF do not completely define the true Bayes' *a posteriori* PDF [22]. Instead of numerically solving FPKE (Equation: 6.1), another useful approximation for state transition PDF and Bayes' *a posteriori* PDF is GCS which is a function of higher order moments of the PDF and Hermite polynomials (see Equation: 5.37 and 5.40) [33],[11]. Culver Filter (CF) approximates Bayes' *a posteriori* PDF as third order GCS [1]. The term third order signifies use of moments and Hermite polynomials up to third order for formulation of state transition and Bayes' *a posteriori* PDFs. In general, differential equations for higher order moments required to formulate GCS (Equation: 5.37 and 5.40) can be derived using Ito differential rule expressed as [23]:

$$E \left[\frac{d\psi[\mathbf{x}(t), t]}{dt} \right] = E \left[\frac{\partial \psi}{\partial t} \right] + E \left[\frac{\partial \psi}{\partial \mathbf{x}} \mathbf{f}[\mathbf{x}(t), t] \right] + \frac{1}{2} \text{Tr} E \left[\mathbf{G}[\mathbf{x}(t), t] \mathbf{Q}(t) \mathbf{G}^T[\mathbf{x}(t), t] \frac{\partial^2 \psi}{\partial^2 \mathbf{x}} \right] \quad (6.10)$$

where, $E[\cdot]$ = expectation operator, $\psi = \psi[\mathbf{x}(t), t]$ and Tr = trace of a matrix.

By selecting the quantity $\psi[\mathbf{x}(t), t]$ as different order moments such as $x_j(t)$, $[x_j(t)x_l(t)]$ and $[x_j(t)x_l(t)x_m(t)]$ etc, where, subscripts $j, l, m \in \{1 \dots d\}$ and d = dimension of the dynamic system, the differential equations of higher order moments up to any order can be generated. However, truncation of moments up to a certain order will be required in order to develop practically feasible filtering algorithms. As discussed in Chapter: 5, the GCS can be considered as a natural extension of Gaussian PDF to approximate arbitrary PDFs. Using third order GCS approximation of state transition and Bayes' *a posteriori* PDFs, CF provides MMSE estimates by solving Bayes' formula (Equation: 6.7) exactly. Much like second order EKF [5] the filter expands the nonlinear advective term of Equation: 6.2 using second order Taylor series expansion. The differential equations for central moments up to third order are derived using Ito differential rule (Equation: 6.10). Using the component wise notation for nonlinear dynamical system expressed in Equation: 6.2, if one considers $\psi = x_i$ where, subscripts $i \in \{1 \dots d\}$ we would get differential equation for mean as:

$$\frac{d\mu_i}{dt} = E[f_i(\mathbf{x}(t), t)] \quad (6.11)$$

Keeping in view differentiability of nonlinear function $\mathbf{f}[(\mathbf{x}(t), t)]$ (of Equation: 6.2), its second order Taylor series expansion around mean μ_i can be expressed as [1]:

$$\frac{dx_i}{dt} = f_i(\boldsymbol{\mu}, t) + F_{ie}(\boldsymbol{\mu}, t)(x_e - \mu_e) + A_{ief}(\boldsymbol{\mu}, t)(x_e - \mu_e)(x_f - \mu_f) + G_{ie}W_e \quad (6.12)$$

where, W_e is the white Gaussian noise, $F_{ie}(\boldsymbol{\mu}, t) = \partial f_i(\boldsymbol{\mu}, t)/\partial x_e$ is a component of Jacobian matrix, $A_{ief}(\boldsymbol{\mu}, t) = \left(\frac{1}{2}\right) \partial^2 f_i(\boldsymbol{\mu}, t)/\partial x_e \partial x_f$ is component of Hessian matrix, tensor subscripts (indices) notations $i, e, f \in \{1 \dots d\}$ assume implicit summation of indices. By taking expectation of Equation: 6.12 and using Ito differential rule the higher order central moments are derived as (see [1] for proof):

$$\frac{d\mu_i(t)}{dt} = f_i(\boldsymbol{\mu}, t) + A_{ief}(\boldsymbol{\mu}, t)P_{ef} \quad (6.13)$$

$$\frac{dP_{ij}(t)}{dt} = 2 \left\{ F_{ie}(\boldsymbol{\mu}, t)P_{je} + A_{ief}(\boldsymbol{\mu}, t)P_{jef}^{(3)} \right\}_s + V_{ij} \quad (6.14)$$

$$\frac{dP_{ijl}^{(3)}(t)}{dt} = 3 \left\{ F_{ie}(\boldsymbol{\mu}, t)P_{jle}^{(3)} + A_{ief}(\boldsymbol{\mu}, t)(P_{je}P_{lf} + P_{jf}P_{le}) \right\}_s \quad (6.15)$$

where, $N\{\cdot\}_s$ is the operation of symmetrising the expression inside the bracket with respect to all subscripts and number N is the number of terms in the expression, for example symmetric terms are expressed as $3\{P_{ij}P_{lm}\}_s = P_{ij}P_{lm} + P_{il}P_{jm} + P_{im}P_{jl}$, V_{ij} are individual components of diffusion matrix of SDE ($\mathbf{G}\mathbf{Q}\mathbf{G}^T$), $P_{ij}(t)$ and $P_{ijl}^{(3)}(t)$ are components of covariance and coskewness tensors respectively. Similar to Equation: 6.12 the quantities on right hand side of these differential equations also assume implicit summation of indices (subscripts). Given some initial estimates for state mean, covariance and coskewness these differential equations (Equation: 6.13 to 6.15) are integrated forward in time to obtain *time update* for these parameters. For *measurement update* the nonlinear measurement function (Equation: 6.5) is also linearized (first order in Taylor series) about the current estimates to approximate measurement likelihood as Gaussian PDF. The state estimates from *time update* are updated on receipt of measurements using following equations known as exact third order

CF measurement updates [1]:

$$\begin{aligned}
 \boldsymbol{v}_k &= \mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_k^-) \\
 \boldsymbol{\Lambda}_k &= \mathbf{H}_k^T \mathbf{R}^{-1} \mathbf{H}_k \\
 \boldsymbol{\Omega}_k &= (\mathbf{P}_k^{-1} + \boldsymbol{\Lambda}_k)^{-1} \\
 \mathbf{d}_k &= \boldsymbol{\Omega}_k \mathbf{H}_k^T \mathbf{R}^{-1} \boldsymbol{v}_k \\
 \boldsymbol{\eta}_k &= \mathbf{P}_k^{-1} \mathbf{d}_k \\
 \boldsymbol{\mathcal{G}}_k &= \boldsymbol{\Omega}_k \mathbf{P}_k^{-1} \\
 \boldsymbol{\mathcal{Q}}_k &= \mathbf{P}_k^{-1} \boldsymbol{\mathcal{G}}_k \\
 \boldsymbol{\varepsilon}_k &= \boldsymbol{\mathcal{Q}}_k - \mathbf{P}_k^{-1} + \frac{1}{3} \boldsymbol{\eta}_k \boldsymbol{\eta}_k^T \\
 \boldsymbol{\mathcal{O}}_k &= \boldsymbol{\mathcal{Q}}_k - \mathbf{P}_k^{-1} + \boldsymbol{\eta}_k \boldsymbol{\eta}_k^T \\
 \\
 \mathcal{V} &= \frac{1}{2} \mathbf{P}_{abc}^{(3)} \eta_a \varepsilon_{bc} \\
 \mathfrak{K} &= \frac{1}{(1 + \mathcal{V})} \\
 \phi_i &= \frac{1}{2} \mathfrak{K} \mathbf{P}_{abc}^{(3)} \mathcal{G}_{ia} \mathcal{O}_{bc} \\
 \varphi_{ij} &= \mathfrak{K} \mathbf{P}_{abc}^{(3)} \eta_a \mathcal{G}_{ib} \mathcal{G}_{jc} \\
 \mathfrak{N}_{ijl} &= \mathfrak{K} \mathbf{P}_{abc}^{(3)} \mathcal{G}_{ia} \mathcal{G}_{jb} \mathcal{G}_{lc}
 \end{aligned} \tag{6.16}$$

where, third order coskewness tensor notations $\mathbf{P}_{abc}^{(3)}$ assumes implicit time subscript “ k ”. Finally the *measurement updated* components of mean, covariance and coskewness of Bayes’ *a posteriori* PDF are [1]:

$$\hat{\mathbf{x}}_l^+ = \hat{\mathbf{x}}_l^- + \mathbf{d}_l + \boldsymbol{\phi}_l \tag{6.17}$$

$$\mathbf{P}_{ij}^+ = \Omega_{ij} + \varphi_{ij} - \phi_i \phi_j \tag{6.18}$$

$$\mathbf{P}_{ijl}^{(3)+} = \mathfrak{N}_{ijl} - \phi_i \varphi_{jl} - \phi_j \varphi_{il} - \phi_l \varphi_{ij} + 2\phi_i \phi_j \phi_l \tag{6.19}$$

6.4 Mixture Culver Filter

The CF is an extension of EKF wherein nonlinear advective term of Equation: 6.2, is linearized up to second order in Taylor series just like second order EKF [5]. However, in CF the assumption of computing only first two central moments and assumption of Gaussian PDF for state transition and Bayes' *a posteriori* PDFs is relaxed as discussed in Section: 6.3. Nevertheless, as discussed in Chapter: 1 (see Figures: 1-3, 1-6) and Chapter: 5 (see Section: 5.3, Figures: 5-4, 5-5) that single GCS of lower order (order ≤ 4) does not provides accurate estimates for mean of true PDF. Moreover, the expansion has regions of negative probabilities. Together these two issues are drawbacks of nonlinear filters based on single GCS of lower order, as the filter; (1) could possibly provide inaccurate mean which is optimal MMSE solution, (2) is susceptible to computational inaccuracies for example, the error covariance matrix output could be non positive definite. In order to improve upon these issues one may extend the order of GCS. However, a higher order GCS is more sensitive to outliers of true PDF [34]. Moreover, analytical solution of Bayes' formula (Equation: 6.7) for higher order GCS approximation of Bayes' *a posteriori* PDF to obtain mean (MMSE solution), and higher order statistics i.e., covariance matrix, coskewness, cokurtosis (fourth order) and fifth order etc would be very complex and hence would be of little use [1]. In Chapter: 5 GCSMM was found as better alternative for improving approximation for arbitrary non-Gaussian PDFs and a new SMC filter was developed known as GCSMPF using such approximation. Therefore, now a new extension is proposed based on third order GCSMM known as MCF. Each GCS component of GCSMM used in MCF is of third order. The basic approach used in MCF is adapted from GSF. However, the algorithms (MCF and GSF) have differences in terms of; (1) use of mixture of GCS instead of mixture of Gaussian for state transition and Bayes' *a posteriori* PDFs, (2) weight updates for each component. Keeping our self in line with this approach one now develops MCF based on the GCSMM. Certain equations are re-expressed (earlier used in the text) in order to facilitate reading. The GCSMM based on third order GCS components can be expressed as:

$$p_{gcsm}(\mathbf{x}, t) = \sum_{g=1}^G \alpha_t^{(g)} \mathcal{N}(\mathbf{x}, \boldsymbol{\mu}_t^{(g)}, \mathbf{P}_t^{(g)}) \left[1 + \sum_{i,j,l} \frac{P_{ijl}^{(3)(g)}}{3!} h_{ijl}(\mathbf{x}_t, \boldsymbol{\mu}_t^{(g)}, \mathbf{P}_t^{(g)}) \right] \quad (6.20)$$

where, t is the time subscript, l is the dimension subscript, $h_{ijl}(\mathbf{x}_t, \boldsymbol{\mu}_t, \mathbf{P}_t)$ are third order multidimensional Hermite polynomial with corresponding input dimensions $i, j, l \in \{1, \dots, d\}$, $P_{ijl}^{(3)}$ is multivariate third central moment, $\alpha_t^{(g)}$ are weight of g^{th} component of GCSMM, G is the total number of GCS components and sum over all input dimensions i, j, l is considered for Hermite polynomials and multivariate moments. Here the fact that each third order central moment $P_{ijl}^{(3)}$ is

equivalent to corresponding third order cumulant $\kappa_{i,j,l}$ [28],[1] is kept into mind to write Equation:. This clarifies the alternative (see Chapter: 5) use of third order cumulants for GCS and GCSMM in Equation: 5.40 and 5.47.

6.4.1 Time Update

Consider Bayes' *a posteriori* PDF expressed in Equation: 6.7 as GCSMM:

$$p(\mathbf{x}_k|\mathbf{y}_k) = \frac{1}{C_k} \sum_{g=1}^G \alpha_{k|k-1}^{(g)} \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)}) \left[1 + \sum_{i,j,l} \frac{P_{ijl}^{(3)(g)}}{3!} h_{ijl}(\mathbf{x}_k, \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)}) \right] p(\mathbf{y}_k|\mathbf{x}_k) \quad (6.21)$$

where,

$k|k-1$ = subscript for state transition PDF, $p(\mathbf{y}_k|\mathbf{x}_k)$ = PDF of measurement conditioned on evolved state and C_k = normalizing constant. Discrete time subscript ($k|k-1$) is used to indicate availability of Bayes' *a posteriori* PDF at discrete measurements instants. In reality state transition PDF for a continuous time dynamical system is obtained by solving FPKE (Equation: 6.1) between the measurements from time t_k to t_{k-1} .

The parameters of $\boldsymbol{\mu}_{k|k-1}$ is the mean, $\mathbf{P}_{k|k-1}$ is covariance and $P_{ijl}^{(3)}$ are coskewness tensor components which can be obtained by numerically integrating Equation: 6.13-6.15. However, in order to compute time update of weights $\alpha_{k|k-1}$ one would adapt the methodology suggested by [49] for a GMM replaced here by a GCSMM. The idea for optimal weight updates for each component of GCSMM is realized by minimizing error between FPKE equation (Equation: 6.1) and time derivative of GCSMM PDF. A continuous time notation (Equation: 6.20) will be used for development of time update of weights. The error in FPKE and time derivative of GCSMM is expressed as:

$$e(\mathbf{x}, t) = \frac{\partial p_{gcsm}(\mathbf{x}, t)}{\partial t} - \mathcal{L}_{FP}(p_{gcsm}(\mathbf{x}, t)) \quad (6.22)$$

where,

$\mathcal{L}_{FP}(\cdot)$ = Fokker-Planck operator [119],[49] is described as:

$$\begin{aligned}\mathcal{L}_{FP}\left(p_{gcsm}(\mathbf{x}, t)\right) &= \sum_{g=1}^G \alpha_t^{(g)} \mathcal{L}_{FP}\left(p_{gcs}^{(g)}\right) \\ &= \mathbf{1}_{FP}^T \boldsymbol{\alpha}_t\end{aligned}\tag{6.23}$$

where, $\boldsymbol{\alpha}_t \in \mathbb{R}^{d \times 1}$ is the vector of weights and the elements of $\mathbf{1}_{FP} \in \mathbb{R}^{d \times 1}$ are given by application of Fokker-Planck operator on individual GCS components:

$$\mathcal{L}_{FP}\left(p_{gcs}^{(g)}\right) = -\frac{\partial^T p_{gcs}^{(g)}}{\partial t} \mathbf{f}(\mathbf{x}, t) - p_{gcs}^{(g)} \text{Tr} \left[\frac{\partial \mathbf{f}(\mathbf{x}, t)}{\partial \mathbf{x}} \right] + \frac{1}{2} \text{Tr} \left[\mathbb{Q} \frac{\partial^2 p_{gcs}^{(g)}}{\partial \mathbf{x} \partial \mathbf{x}^T} \right]\tag{6.24}$$

The first term on right of Equation: 6.22 is obtained by taking total derivative expressed as:

$$\frac{\partial p_{gcsm}(\mathbf{x}, t)}{\partial t} = \sum_{g=1}^G \left(\dot{\alpha}_t^{(g)} p_{gcs}^{(g)} + \alpha_t^{(g)} \frac{\partial^T p_{gcs}^{(g)}}{\partial \boldsymbol{\mu}_t^{(g)}} \dot{\boldsymbol{\mu}}_t^{(g)} + \alpha_t^{(g)} \text{Tr} \left[\frac{\partial p_{gcs}^{(g)}}{\partial \mathbf{P}_t^{(g)}} \dot{\mathbf{P}}_t^{(g)} \right] + \alpha_t^{(g)} \frac{\partial p_{gcs}^{(g)}}{\partial P_{ijl}^{(3)(g)}} \dot{P}_{ijl}^{(3)(g)} \right)\tag{6.25}$$

where,

Tr = trace and the last term in above equation (Equation: 6.25) implies summation of derivatives over all indices (i, j, l) obtained as:

$$\alpha_t^{(g)} \sum_{i,j,l} \frac{\partial p_{gcs}^{(g)}}{\partial P_{ijl}^{(3)(g)}} \dot{P}_{ijl}^{(3)(g)}\tag{6.26}$$

The total derivative of the moments $\dot{\boldsymbol{\mu}}_t^{(g)}$, $\dot{\mathbf{P}}_t^{(g)}$ and $\dot{P}_{ijl}^{(3)(g)}$ for each GCS component is given in Equation: 6.13-6.15 and the derivative of weights, $\dot{\alpha}_t^{(g)}$ is obtained by time discretization using the

first forward difference [120]:

$$\dot{\alpha}_t^{(g)} = \frac{1}{\Delta t} (\alpha_{t'}^{(g)} - \alpha_t^{(g)}) \quad (6.27)$$

where,

$$t' = t + \Delta t$$

Now by substituting Equation: 6.27 into Equation: 6.25 one may rewrite total time derivative of GCSMM as:

$$\begin{aligned} \frac{\partial p_{gcsm}(\mathbf{x}, t)}{\partial t} &= \sum_{g=1}^G \frac{1}{\Delta t} p_{gcs}^{(g)} \alpha_{t'}^{(g)} \\ &+ \underbrace{\sum_{g=1}^G \left(\frac{\partial^T p_{gcs}^{(g)}}{\partial \boldsymbol{\mu}_t^{(g)}} \dot{\boldsymbol{\mu}}_t^{(g)} + \text{Tr} \left[\frac{\partial p_{gcs}^{(g)}}{\partial \mathbf{P}_t^{(g)}} \dot{\mathbf{P}}_t^{(g)} \right] + \frac{\partial p_{gcs}^{(g)}}{\partial \mathbf{P}_{ijl}^{(3)(g)}} \dot{\mathbf{P}}_{ijl}^{(3)(g)} - \frac{1}{\Delta t} p_{gcs}^{(g)} \right) \alpha_t^{(g)}}_{\mathbf{m}_{TD}^{(g)}} \end{aligned} \quad (6.28)$$

$$= \frac{1}{\Delta t} \mathbf{p}_{gcs}^T \boldsymbol{\alpha}_{t'} + \mathbf{m}_{TD}^T \boldsymbol{\alpha}_t \quad (6.29)$$

where, $\boldsymbol{\alpha}_{t'} \in \mathbb{R}^{d \times 1}$ is the vector of new weights which are being found out, $\mathbf{p}_{gcs} \in \mathbb{R}^{d \times 1}$ is the vector of GCS components and the elements of $\mathbf{m}_{TD} \in \mathbb{R}^{d \times 1}$ are expressed in Equation: 6.28. Now by substituting Equations: 6.23 and 6.29 into Equation: 6.22 one would get FPKE error as:

$$e(t, \mathbf{x}) = \frac{1}{\Delta t} \mathbf{p}_{gcs}^T \boldsymbol{\alpha}_{t'} + (\mathbf{m}_{TD} - \mathbf{1}_{FP})^T \boldsymbol{\alpha}_t \quad (6.30)$$

Furthermore, analytical expressions for different derivatives used in Equation: 6.28 can be conveniently expressed in component wise tensor notation as:

$$\begin{aligned} \frac{\partial p_{gcs}}{\partial \mu_a} &= \left[\mathbf{P}_{au}^{-1} (x_u - \mu_u) \right. \\ &+ \frac{1}{3!} \mathbf{P}_{ijl}^{(3)} (\mathbf{P}_{au}^{-1} (x_u - \mu_u) h_{ijl} - \mathbf{P}_{ia}^{-1} h_j h_l - \mathbf{P}_{ja}^{-1} h_i h_l - \mathbf{P}_{la}^{-1} h_i h_j + \mathbf{P}_{ia}^{-1} \mathbf{P}_{jl}^{-1} \\ &\left. + \mathbf{P}_{ja}^{-1} \mathbf{P}_{il}^{-1} + \mathbf{P}_{la}^{-1} \mathbf{P}_{ij}^{-1}) \right] p_g \end{aligned} \quad (6.31)$$

$$\frac{\partial p_{gcs}}{\partial P_{mn}} = \left[\frac{1}{2} (P_{mr}^{-1} P_{ns}^{-1} (x_r - \mu_r)(x_s - \mu_s) - P_{mn}^{-1}) \right. \quad (6.32)$$

$$\begin{aligned} & + \frac{1}{3!} P_{ijl}^{(3)} \left(\frac{1}{2} (P_{mr}^{-1} P_{ns}^{-1} (x_r - \mu_r)(x_s - \mu_s) - P_{mn}^{-1}) h_{ijl} \right. \\ & - P_{im}^{-1} P_{nr}^{-1} P_{js}^{-1} P_{lv}^{-1} (x_r - \mu_r)(x_s - \mu_s)(x_v - \mu_v) \\ & - P_{jm}^{-1} P_{ns}^{-1} P_{ir}^{-1} P_{lv}^{-1} (x_r - \mu_r)(x_s - \mu_s)(x_v - \mu_v) \\ & - P_{lm}^{-1} P_{nv}^{-1} P_{ir}^{-1} P_{js}^{-1} (x_r - \mu_r)(x_s - \mu_s)(x_v - \mu_v) + P_{im}^{-1} P_{jl}^{-1} P_{nr}^{-1} (x_r - \mu_r) \\ & + P_{jm}^{-1} P_{nl}^{-1} P_{ir}^{-1} (x_r - \mu_r) + P_{il}^{-1} P_{jm}^{-1} P_{ns}^{-1} (x_s - \mu_s) + P_{im}^{-1} P_{nl}^{-1} P_{js}^{-1} (x_s - \mu_s) \\ & \left. \left. + P_{ij}^{-1} P_{lm}^{-1} P_{nt}^{-1} (x_v - \mu_v) + P_{im}^{-1} P_{nj}^{-1} P_{lv}^{-1} (x_v - \mu_v) \right) \right] p_g \end{aligned}$$

$$\frac{\partial p_{gcs}}{\partial x_a} = \left[-P_{au}^{-1} (x_u - \mu_u) \right. \quad (6.33)$$

$$\begin{aligned} & + \frac{1}{3!} P_{ijl}^{(3)} (-P_{au}^{-1} (x_u - \mu_u) h_{ijl} \\ & + P_{ia}^{-1} P_{js}^{-1} P_{lv}^{-1} (x_s - \mu_s)(x_v - \mu_v) + P_{ja}^{-1} P_{ir}^{-1} P_{lv}^{-1} (x_r - \mu_r)(x_v - \mu_v) \\ & + P_{la}^{-1} P_{ir}^{-1} P_{js}^{-1} (x_r - \mu_r)(x_s - \mu_s) - P_{ia}^{-1} P_{jl}^{-1} - P_{ja}^{-1} P_{il}^{-1} - P_{la}^{-1} P_{ij}^{-1}) \left] p_g \end{aligned}$$

$$\frac{\partial^2 p_{gcs}}{\partial x_m \partial x_n} = \left[(P_{mr}^{-1} P_{ns}^{-1} (x_r - \mu_r)(x_s - \mu_s) - P_{mn}^{-1}) \right. \quad (6.34)$$

$$\begin{aligned} & + \frac{1}{3!} P_{ijl}^{(3)} \left(\frac{1}{2} (P_{mr}^{-1} P_{ns}^{-1} (x_r - \mu_r)(x_s - \mu_s) - P_{mn}^{-1}) h_{ijl} \right. \\ & + P_{im}^{-1} P_{jn}^{-1} P_{lv}^{-1} (x_v - \mu_v) + P_{im}^{-1} P_{ln}^{-1} P_{js}^{-1} (x_s - \mu_s) + P_{jm}^{-1} P_{in}^{-1} P_{lv}^{-1} (x_v - \mu_v) \\ & + P_{jm}^{-1} P_{ln}^{-1} P_{ir}^{-1} (x_r - \mu_r) + P_{lm}^{-1} P_{in}^{-1} P_{js}^{-1} (x_s - \mu_s) \\ & \left. \left. + P_{lm}^{-1} P_{jn}^{-1} P_{ir}^{-1} (x_r - \mu_r) \right) \right] p_g \end{aligned}$$

$$\frac{\partial p_{gcs}}{\partial P_{ijl}^{(3)}} = \left[\frac{1}{3!} h_{ijl} \right] p_g \quad (6.35)$$

where, component wise tensor notation used in above expressions utilizes implicit summation of indices and p_g denotes Gaussian PDF. For other notations used in above expressions see Section: 5.3 (Equation: 5.42-5.43). Now by propagating the mean $\mu_t^{(g)}$, covariance $\mathbf{P}_t^{(g)}$ and $P_{ijl}^{(3)(g)}$ of individual

GCS component using Equation: 6.13-6.15 one seeks to obtain new weights by minimizing the error in FPKE over a selected volume of state space [120]:

$$\begin{aligned} \min_{\alpha_{t'}^{(g)}} & \frac{1}{2} \int_V e^2(t, \mathbf{x}) d\mathbf{x} \\ \text{s.t.} & \sum_{g=1}^G \alpha_{t'}^{(g)} = 1 \\ & \alpha_{t'}^{(g)} \geq 0, \quad g = 1 \dots G \end{aligned} \quad (6.36)$$

The aforementioned problem can be written as a quadratic programming problem for which efficient solvers are available in programming languages such as Matlab i.e., quadprog [120]:

$$\begin{aligned} \min_{\alpha_{t'}^{(g)}} & \frac{1}{2} \alpha_{t'}^T \mathbf{M}_c \alpha_t + \alpha_{t'}^T \mathbf{N}_c \alpha_t + (\alpha_{t'} - \alpha_t)^T (\alpha_{t'} - \alpha_t) \\ \text{s.t.} & \mathbf{1}_{d \times 1}^T \alpha_{t'} = 1 \\ & \alpha_{t'} \geq \mathbf{0}_{d \times 1} \end{aligned} \quad (6.37)$$

where, $\mathbf{1}_{d \times 1} \in \mathbb{R}^{d \times 1}$ is a vector of ones, $\mathbf{0}_{d \times 1} \in \mathbb{R}^{d \times 1}$ is a vector of zeros and the matrices $\mathbf{M}_c \in \mathbb{R}^{d \times d}$ and $\mathbf{N}_c \in \mathbb{R}^{d \times d}$ are given by:

$$\begin{aligned} \mathbf{M}_c &= \frac{1}{\Delta t^2} \int_V \mathbf{p}_{gcs} \mathbf{p}_{gcs}^T d\mathbf{x} \\ \mathbf{N}_c &= \frac{1}{\Delta t} \int_V \mathbf{p}_{gcs} (\mathbf{m}_{TD} - \mathbf{1}_{FP})^T d\mathbf{x} \end{aligned} \quad (6.38)$$

Analytical solutions were found out for above integrals and presented in Appendix-G. To author's knowledge the adaptation of FPKE error feedback methodology [120] for GCSMM using analytical or numerical methods is new and has not appeared anywhere in estimation and filtering literature.

6.4.2 Measurement Update

Using the time updated GCSMM along with new weights for each GCS component $\alpha_{k|k-1}^{(g)}$, one now consider treatment of Bayes' *a posteriori* PDF (Equation: 6.21) for MMSE solution of our filtering problem. Firstly the normalization constant C_k in Equation: 6.21 can be obtained as:

$$C_k \triangleq p(\mathbf{y}_k | \mathbf{y}_{k-1}) \quad (6.39)$$

$$= \left(\int_{-\infty}^{+\infty} \sum_{g=1}^G \alpha_{k|k-1}^{(g)} \mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)}) \left[1 + \sum_{i,j,l} \frac{P_{ijl}^{(3)(g)}}{3!} h_{ijl}(\mathbf{x}_k, \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)}) \right] p(\mathbf{y}_k | \mathbf{x}_k) d\mathbf{x}_k \right)$$

Each “ g ” GCS component inside integral of Equation: 6.39 can be written as:

$$\int_{-\infty}^{+\infty} \left\{ \mathcal{N}(\mathbf{x}_k, \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)}) \left[1 + \sum_{i,j,l} \frac{P_{ijl}^{(3)(g)}}{3!} h_{ijl}(\mathbf{x}_k, \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)}) \right] p(\mathbf{y}_k | \mathbf{x}_k) \right\} d\mathbf{x}_k \quad (6.40)$$

By linearizing the measurement function $\mathbf{h}(\cdot)$ (Equation: 6.5) using first order Taylor series expansion, around predicted estimates $\boldsymbol{\mu}_{k|k-1}^{(g)}$ one may approximate $p(\mathbf{y}_k | \mathbf{x}_k)$ as a multidimensional Gaussian PDF [1]:

$$p(\mathbf{y}_k | \mathbf{x}_k) = \frac{1}{\sqrt{|2\pi\mathbf{R}|}} \exp \left[-\frac{1}{2} \left(\mathbf{y}_k - \mathbf{h}(\boldsymbol{\mu}_{k|k-1}^{(g)}) - \mathbf{H}_k (\mathbf{x}_k - \boldsymbol{\mu}_{k|k-1}^{(g)}) \right)^T \mathbf{R}^{-1} \left(\mathbf{y}_k - \mathbf{h}(\boldsymbol{\mu}_{k|k-1}^{(g)}) - \mathbf{H}_k (\mathbf{x}_k - \boldsymbol{\mu}_{k|k-1}^{(g)}) \right) \right] \quad (6.41)$$

where, $\mathbf{H}_k = (\partial \mathbf{h} / \partial \mathbf{x}_k) |_{\mathbf{x}_k = \boldsymbol{\mu}_{k|k-1}^{(g)}}$

By substituting Equation: 6.41 in Equation: 6.40 and solving the integral would give (see reference [1] for proof):

$$\begin{aligned}
& D_k^{(g)} \\
& = \frac{\exp \left\{ -\frac{1}{2} \left(\mathbf{y}_k - \mathbf{h} \left(\boldsymbol{\mu}_{k|k-1}^{(g)} \right) \right)^T \left(\mathbf{R}^{-1} - \mathbf{R}^{-1} \mathbf{H}_{k|k-1}^{(g)} \boldsymbol{\Omega}_k^{(g)} \mathbf{H}_{k|k-1}^{T(g)} \mathbf{R}^{-1} \right) \left(\mathbf{y}_k - \mathbf{h} \left(\boldsymbol{\mu}_{k|k-1}^{(g)} \right) \right) \right\}}{\sqrt{\left| \mathbf{I} + \mathbf{H}_{k|k-1}^{T(g)} \mathbf{R}^{-1} \mathbf{H}_{k|k-1}^{(g)} \mathbf{P}_{k|k-1}^{(g)} \right|}} \\
& \quad \times \left(1 + \mathcal{V}_k^{(g)} \right)
\end{aligned} \tag{6.42}$$

where, \mathbf{I} denotes identity matrix. Therefore, by integrating each term inside integral of Equation: 6.40 would yield the denominator C_k as:

$$C_k = \sum_{g=1}^G \alpha_{k|k-1}^{(g)} D_k^{(g)} \tag{6.43}$$

Now mean, covariance and coskewness tensor of Bayes' *a posteriori* PDF (Equation: 6.21) can be calculated using following integrals:

$$\begin{aligned}
\mathbf{x}_k^+ &= \frac{1}{C_k} \int_{-\infty}^{+\infty} p(\mathbf{x}_k | \mathbf{y}_k) \mathbf{x}_k d\mathbf{x}_k \\
\mathbf{P}_k^+ &= \frac{1}{C_k} \int_{-\infty}^{+\infty} p(\mathbf{x}_k | \mathbf{y}_k) (\mathbf{x}_k - \mathbf{x}_k^+) (\mathbf{x}_k - \mathbf{x}_k^+)^T d\mathbf{x}_k \\
\mathbf{P}_k^{+(3)} &= \frac{1}{C_k} \int_{-\infty}^{+\infty} p(\mathbf{x}_k | \mathbf{y}_k) (\mathbf{x}_k - \mathbf{x}_k^+) (\mathbf{x}_k - \mathbf{x}_k^+)^T \otimes (\mathbf{x}_k - \mathbf{x}_k^+)^T d\mathbf{x}_k
\end{aligned} \tag{6.44}$$

Firstly one compute mean \mathbf{x}_k^+ by rewriting first of Equation: 6.44 as:

$$\begin{aligned}
\mathbf{x}_k^+ &= \frac{1}{\sum_{g=1}^G \alpha_{k|k-1}^{(g)} D_k^{(g)}} \int_{-\infty}^{+\infty} \sum_{g=1}^G \alpha_{k|k-1}^{(g)} \mathcal{N} \left(\mathbf{x}_k; \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)} \right) \left[1 \right. \\
& \quad \left. + \sum_{i,j,l} \frac{\mathbf{P}_{ijl}^{(g)}}{3!} h_{ijl} \left(\mathbf{x}_k, \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)} \right) \right] p(\mathbf{y}_k | \mathbf{x}_k) \mathbf{x}_k d\mathbf{x}_k
\end{aligned} \tag{6.45}$$

The integral in Equation: 6.45 can be solved by treating each GCS component individually as:

$$\mathbf{x}_k^{+(g)} = \frac{1}{D_k^{(g)}} \int_{-\infty}^{+\infty} p_{gcs}^{(g)}(\mathbf{x}_{k|k-1} | \mathbf{y}_{k-1}) p(\mathbf{y}_k | \mathbf{x}_k) \mathbf{x}_k d\mathbf{x}_k \quad (6.46)$$

where,

$$p_{gcs}^{(g)}(\mathbf{x}_{k|k-1} | \mathbf{y}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)}) \left[1 + \sum_{i,j,l} \frac{P_{ijl}^{(g)}}{3!} h_{ijl}(\mathbf{x}_k, \boldsymbol{\mu}_{k|k-1}^{(g)}, \mathbf{P}_{k|k-1}^{(g)}) \right] \quad (6.47)$$

Similar treatment for computation of covariance and coskewness tensor for each GCS component yields:

$$\begin{aligned} \mathbf{P}_k^{+(g)} &= \frac{1}{D_k^{(g)}} \int_{-\infty}^{+\infty} p_{gcs}^{(g)}(\mathbf{x}_k | \mathbf{y}_k) (\mathbf{x}_k - \mathbf{x}_k^{+(g)}) (\mathbf{x}_k - \mathbf{x}_k^{+(g)})^T d\mathbf{x}_k \\ \mathbf{P}_k^{+(3)(g)} &= \frac{1}{D_k^{(g)}} \int_{-\infty}^{+\infty} p_{gcs}^{(g)}(\mathbf{x}_k | \mathbf{y}_k) (\mathbf{x}_k - \mathbf{x}_k^{+(g)}) (\mathbf{x}_k - \mathbf{x}_k^{+(g)})^T \otimes (\mathbf{x}_k - \mathbf{x}_k^{+(g)})^T d\mathbf{x}_k \end{aligned} \quad (6.48)$$

By using CF measurement update equations (Equation: 6.17 to 6.19) the solution for each of above integral (Equation: 6.46 and 6.48) is given as:

$$\hat{\mathbf{x}}_i^{+(g)} = \hat{\mathbf{x}}_i^{-(g)} + \mathbf{d}_i^{(g)} + \phi_i^{(g)} \quad (6.49)$$

$$P_{ij}^{+(g)} = \Omega_{ij}^{(g)} + \varphi_{ij}^{(g)} - \phi_i^{(g)} \phi_j^{(g)}$$

$$P_{ijl}^{+(3)(g)} = \mathfrak{N}_{ijl}^{(g)} - \phi_i^{(g)} \varphi_{jl}^{(g)} - \phi_j^{(g)} \varphi_{il}^{(g)} - \phi_l^{(g)} \varphi_{ij}^{(g)} + 2\phi_i^{(g)} \phi_j^{(g)} \phi_l^{(g)}$$

The weights could be conveniently updated using zero moment $D_k^{(g)}$ (Equation: 6.42) of each GCS

component of the Bayes' *a posteriori* PDF:

$$\alpha_k^{(g)} = \frac{\alpha_{k|k-1}^{(g)} D_k^{(g)}}{\sum_{g=1}^G \alpha_{k|k-1}^{(g)} D_k^{(g)}} \quad (6.50)$$

MCF can be initialized using EM algorithm [107],[108] already explained in Chapter: 5. In order to simplify computation initial PDF can be assumed as GMM. A RR step (see Section: 5.2.4) in MCF algorithm after weight update is added to produce children of GCS components having more significant weights and discarding components having insignificant weights. The resampling strategy for mixands in MCF and GSF (later used for comparison) is adapted from [64]. Thus, the effective size of weight G_E could be expressed as:

$$G_E = \frac{1}{\sum_{g=1}^G \left(\alpha_k^{(g)}\right)^2} \quad (6.51)$$

If $G_E < G_T$ where G_T is required (threshold) size of weights we would perform RR step.

Time update of weights for each GCS component in MCF described in Section: 6.4.1 could become quite extensive for higher dimensional systems. Therefore, one can simplify the algorithm by keeping the weights constant between the measurements. This is essentially the same methodology used in traditional GSF. In sequel a comparison for benefits of full MCF algorithm (complete with time update of weights as described in Section: 6.4.1) and simplified MCF wherein weight of each GCS mixand is kept constant between the measurements is being done. Thus one is now able to furnish a computational algorithm for a MCF as shown in Table: 6.3. It is well documented in the estimation theory literature that with just a minor change in mechanization ahead of standard EKF (as in case of GSF) implementation can result in a significant improvement in EKF performance [121]. Therefore, one expects minor changes in single GCS filter to obtain significant filtering performance. The comparison of MCF with single GCS filter (CF) [1], GSF [44] and EKF [5][10] has been carried out. The new filter has shown improvement over other methods especially under uncertain initial conditions and sparse data availability.

1. Initial estimates / higher order statistics and noise statistics:

$$\hat{\mathbf{x}}_0^+, \mathbf{P}_0^+, \mathbf{P}_0^{+(3)}, \quad p(\mathbf{W}) = \mathcal{N}(0, \mathbf{G}\mathbf{Q}\mathbf{G}^T)$$

2. Perform EM to obtain GCSMM from (1)

3. Time Update - State Propagation: For $g = 1 \dots G$

$$\hat{\mathbf{x}}_{i(k)}^{-(g)} = \int_{t_{k-1}}^{t_k} \frac{d\hat{\mathbf{x}}_i^{(g)}(t)}{dt} = \mathbf{f}_i^{(g)}(\hat{\mathbf{x}}, t) + \mathbf{A}_{ief}^{(g)}(\hat{\mathbf{x}}, t) \mathbf{P}_{ef}^{(g)} dt$$

$$\mathbf{P}_{ij(k)}^{-(g)} = \int_{t_{k-1}}^{t_k} \frac{d\mathbf{P}_{ij}^{(g)}(t)}{dt} = 2 \left\{ \mathbf{F}_{ie}^{(g)}(\hat{\mathbf{x}}, t) \mathbf{P}_{je}^{(g)} + \mathbf{A}_{ief}^{(g)}(\hat{\mathbf{x}}, t) \mathbf{P}_{jef}^{(3)(g)} \right\}_s + \mathbf{V}_{ij} dt$$

$$\mathbf{P}_{ijl(k)}^{-(3)(g)} = \int_{t_{k-1}}^{t_k} \frac{d\mathbf{P}_{ijl}^{(3)(g)}(t)}{dt} = 3 \left\{ \mathbf{F}_{ie}^{(g)}(\hat{\mathbf{x}}, t) \mathbf{P}_{jle}^{(3)(g)} + \mathbf{A}_{ief}^{(g)}(\hat{\mathbf{x}}, t) (\mathbf{P}_{je}^{(g)} \mathbf{P}_{lf}^{(g)} + \mathbf{P}_{jf}^{(g)} \mathbf{P}_{le}^{(g)}) \right\}_s dt$$

4. Time update of GCSMM weights $\alpha_{k|k-1}$ (Optional see text for remarks)

5. Measurement Update: For $g = 1 \dots G$

$$\begin{aligned} \mathbf{v}_k^{(g)} &= \mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_k^{-(g)}), \quad \mathbf{\Lambda}_k^{(g)} = \mathbf{H}_k^T \mathbf{R}^{-1} \mathbf{H}_k^{(g)}, \quad \mathbf{\Omega}_k^{(g)} = (\mathbf{P}_k^{-1(g)} + \mathbf{\Lambda}_k^{(g)})^{-1} \\ \mathbf{d}_k^{(g)} &= \mathbf{\Omega}_k^{(g)} \mathbf{H}_k^T \mathbf{R}^{-1} \mathbf{v}_k^{(g)}, \quad \mathbf{\eta}_k^{(g)} = \mathbf{P}_k^{-1(g)} \mathbf{d}_k^{(g)}, \quad \mathbf{g}_k^{(g)} = \mathbf{\Omega}_k^{(g)} \mathbf{P}_k^{-1(g)}, \quad \mathbf{q}_k^{(g)} = \mathbf{P}_k^{-1(g)} \mathbf{g}_k^{(g)}, \\ \mathbf{\varepsilon}_k^{(g)} &= \mathbf{q}_k^{(g)} - \mathbf{P}_k^{-1(g)} + \frac{1}{3} \mathbf{\eta}_k^{(g)} \mathbf{\eta}_k^{T(g)}, \quad \mathbf{o}_k^{(g)} = \mathbf{q}_k^{(g)} - \mathbf{P}_k^{-1(g)} + \mathbf{\eta}_k^{(g)} \mathbf{\eta}_k^{T(g)} \end{aligned}$$

6. Tensor Notations (time subscript “ k ” is removed for clarity)

$$\mathcal{V}^{(g)} = \frac{1}{2} \mathbf{P}_{abc}^{(3)(g)} \eta_a^{(g)} \varepsilon_{bc}^{(g)}, \quad \mathfrak{K}^{(g)} = \frac{1}{(1 + \mathcal{V}^{(g)})}, \quad \phi_i^{(g)} = \frac{1}{2} \mathfrak{K}^{(g)} \mathbf{P}_{abc}^{(3)(g)} \mathcal{G}_{ia}^{(g)} \mathcal{O}_{bc}^{(g)}$$

$$\phi_{ij}^{(g)} = \mathfrak{K}^{(g)} \mathbf{P}_{abc}^{(g)(3)} \eta_a^{(g)} \mathcal{G}_{ib}^{(g)} \mathcal{G}_{jc}^{(g)}, \quad \mathfrak{N}_{ijl}^{(g)} = \mathfrak{K}^{(g)} \mathbf{P}_{abc}^{(3)(g)} \mathcal{G}_{ia}^{(g)} \mathcal{G}_{jb}^{(g)} \mathcal{G}_{lc}^{(g)}$$

$$\hat{\mathbf{x}}_i^{+(g)} = \hat{\mathbf{x}}_i^{-(g)} + \mathbf{d}_i^{(g)} + \phi_i^{(g)}, \mathbf{P}_{ij}^{+(g)} = \mathbf{\Omega}_{ij}^{(g)} + \phi_{ij}^{(g)} - \phi_i^{(g)} \phi_j^{(g)}$$

$$\mathbf{P}_{ijl}^{+(3)(g)} = \mathfrak{N}_{ijl}^{(g)} - \phi_i^{(g)} \phi_{jl}^{(g)} - \phi_j^{(g)} \phi_{il}^{(g)} - \phi_l^{(g)} \phi_{ij}^{(g)} + 2\phi_i^{(g)} \phi_j^{(g)} \phi_l^{(g)}$$

7. Weight Updates:
$$\alpha_k^{(g)} = \frac{\alpha_{k-1}^{(g)} \mathbf{D}_{k-1}^{(g)}}{\sum_{g=1}^G \alpha_{k-1}^{(g)} \mathbf{D}_{k-1}^{(g)}}$$

Optional: RR Step: $G_E = \frac{1}{\sum_{g=1}^G (\alpha_k^{(g)})^2}$, where $G_E < G_T$ is prescribed threshold criteria

8. Inference: The conditional mean state estimate $\hat{\mathbf{x}}_k^+$ and Covariance \mathbf{P}_k^+ can be estimated by:

$$\hat{\mathbf{x}}_k^+ = \sum_{g=1}^G \alpha_k^{(g)} \hat{\mathbf{x}}_k^{+(g)}, \mathbf{P}_k^+ = \sum_{g=1}^G \alpha_k^{(g)} \left(\mathbf{P}_k^{+(g)} + (\hat{\mathbf{x}}_k^{+(g)} - \hat{\mathbf{x}}_k^+) (\hat{\mathbf{x}}_k^{+(g)} - \hat{\mathbf{x}}_k^+)^T \right)$$

Table 6-1: Mixture Culver Filter

6.5 Orbit Determination using Radar Measurements

In this section algorithms for *Kalman* and *Culver* filter frameworks would now be implemented for satellite OD using ground based radars. The equations of motion for *true model* used in this experiment are given as:

$$\dot{\mathbf{r}} = \mathbf{v} \quad (6.52)$$

$$\dot{\mathbf{v}} = -\frac{\mu_E}{r^3} \mathbf{r} + \mathbf{a}_G + \mathbf{a}_D \quad (6.53)$$

$$\mathbf{a}_D = \frac{1}{2} \rho \frac{C_D A}{m} v^2 \frac{\mathbf{v}}{|\mathbf{v}|} \quad (6.54)$$

where,

$\mathbf{r} = [X, Y, Z]^T$, $\mathbf{v} = [\dot{X}, \dot{Y}, \dot{Z}]^T$ are position and velocity of a satellite in ECI coordinates,

\mathbf{a}_G = perturbation acceleration due to zonal gravitational harmonic up to J_4 , and

\mathbf{a}_D = atmospheric drag acceleration.

The parameters used in Equation: 6.54 are assumed as:

mass ($m = 100$ kg), cross-sectional area ($A = 1$ m²),

atmospheric density ($\rho = 2.043 \times 10^{-14}$) and

drag coefficient ($C_D = 2.2$) [12].

Given some specific initial conditions $\mathbf{x}_0 = [\mathbf{r}_0 \ \mathbf{v}_0]^T$ these equations (Equations: 6.52 to 6.54) are integrated using numerical method such as Adams-Bashforth-Moulton PECE solver `ode113` (with adaptive time step) of Matlab to get time history of position and velocity in ECI reference frame termed here as *true trajectory*. The *true trajectory* is being measured by a radar system fixed at some location on Earth. Reader is referred to Section: 5.7.3 for details on measurement model, radar site location and other measurement parameters i.e., noise variances.

The *filter model* used in this experiment is identical to Equations: 5.92 (2 body dynamics perturbed by J_2 only). However, in *continuous-discrete filtering* one would integrate differential equations for mean, covariance and coskewness tensor (with variable time step) for *time update*. An important computational aspect of higher order filters like CF is the increase in the number of differential equations vis-à-vis increase in the dimension of the system.

For an OD problem the number of differential equations for EKF and CF are tabulated in Table: 6-2. The comparison clearly indicates the computational intensiveness of higher order filters especially once MCF and GSF are being *time updated*.

Type of Filter	Number of differential equations required to compute each given moment			Total number of differential equations
	1 st order moment	2 nd order moment	3 rd order moment	
EKF	6	21	-	27
CF	6	21	56	83

Table 6-2: Comparison of number of first order differential equations for time updates in EKF and CF.

6.5.1 State Uncertainty and Sparse Measurements

In this section one would carry out OD for the satellite in a LEO orbit under sparse measurements i.e., the measurements are available for approximately 4% for orbital period. The reason for this selection is already described in Section: 5.7.3. Moreover, an analysis of the filtering performance under highly uncertain initial conditions for estimates of state, covariance and coskewness tensors is done. Space object initial estimates could be extremely uncertain especially in case of a sparsely tracked object. Therefore, one would now observe filtering performance with increased uncertainty in position variances as 10^7 m^2 and velocity variance $50000 \text{ m}^2 \cdot \text{s}^{-2}$. The position and velocity deviation of our initial estimate ($\hat{\mathbf{x}}_0$) from true initial state (\mathbf{x}_0) is 3162 m and $223 \text{ m} \cdot \text{s}^{-1}$ respectively. This could be considered as significant initial deviation. The initial conditions for the satellite used to generate *true trajectory* are given as (adapted from reference [59]):

$$\hat{\mathbf{P}}_0 = \text{diag}([10^7 \ 10^7 \ 10^7 \ 50000 \ 50000 \ 50000]), \hat{\mathbf{P}}_0^{(3)} = \hat{\mathbf{P}}_{ijl}^{(3)} = 0 \quad (6.55)$$

$$\hat{\mathbf{x}}_0 = [\mathbf{x}_0] + \left[\sqrt{10^7} \ \sqrt{10^7} \ \sqrt{10^7} \ \sqrt{50000} \ \sqrt{50000} \ \sqrt{50000} \right]^T$$

The process noise is selected as $\mathbf{G}\mathbf{Q}\mathbf{G}^T = \text{diag}([0 \ 0 \ 0 \ 10^{-8} \ 10^{-8} \ 10^{-8}]) \text{ m}^2 \cdot \text{s}^{-2}$ due to zonal harmonic J_3, J_4 (order of $0.07 - 0.09 \times 10^{-3} \text{ m} \cdot \text{s}^{-2}$) [14] and atmospheric drag term (see Table: 5-1 for details) used for true model shown in Equation: 6.52 to 6.54. The error criteria of IE and RMSE are being used to gauge filtering performance. Firstly, one provides the measurement data availability of 1 Hz. Due to this high frequency of measurement availability the time update (Step-4, Table: 6-1) is avoided in our first simulation. The time history of IE in ECI coordinates are given in Figure: 6.3 and 6.4 shows that the estimates of CF, MCF, GSF and EKF are close to each other. The convergence to lower errors of MCF is comparatively better than other three filters. The MCF and GSF are both

being propagated using two GCSMM and GMM components respectively. The error plots of these figures (Figure: 6.3 and 6.4) are obtained after averaging 100 MC runs for each filter.

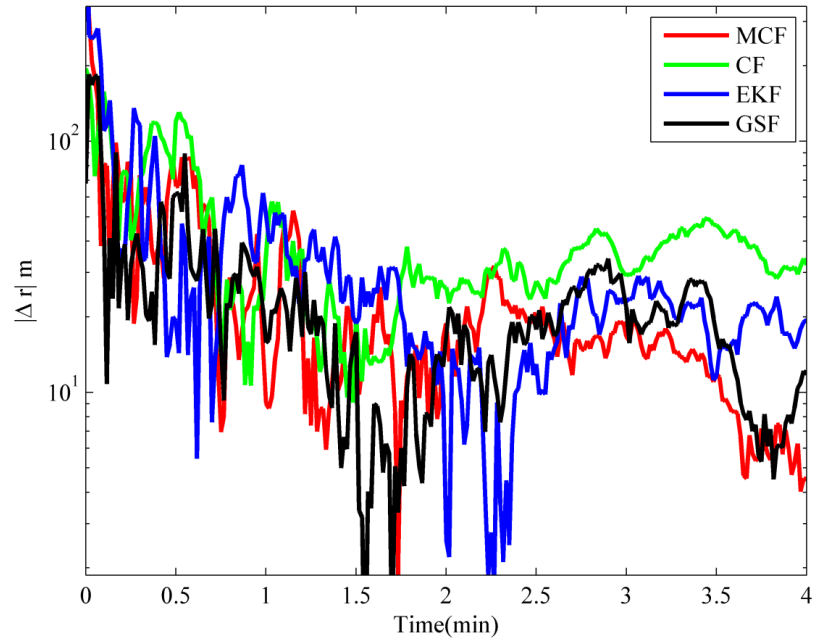


Figure 6-3: Time history of absolute position errors (Δ) in ECI coordinates (shown in log scale) for filters with initial conditions of Equation: 6.55 and measurement availability is 1 Hz.

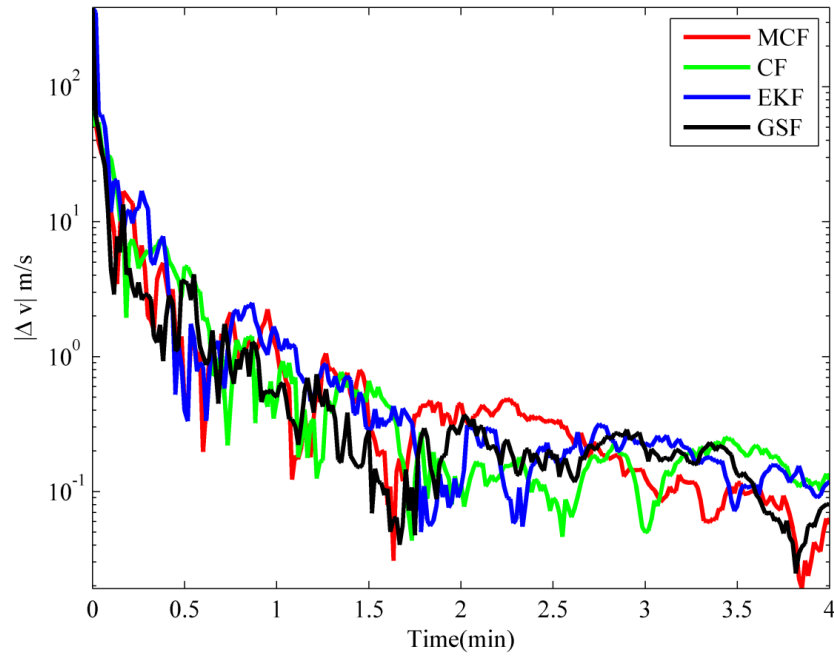


Figure 6-4: Time history of absolute velocity errors (Δ) in ECI coordinates (shown in log scale) for filters with initial conditions of Equation: 6.55 and measurement availability of 1 Hz.

One of the drawbacks of filters based on mixture PDFs is the suboptimal *time update* of mixand weights when there are fewer or no measurements. In this situation the weights would remain constant until a measurement is received. This could possibly produce inferior estimates for filters based on mixture models. One would now incorporate optimal time update of weights (described in Section: 6.4.1) in MCF algorithm to compare filters for 0.033 Hz measurement availability (see Figures: 6.5) and also consider filtering performance over period of time once no observation is available. This frequency of measurement would require optimal time update of weights (Step-4, Table: 6.1). The time history of IE is shown in Figure: 6.5. The error curves of pair, (1) MCF and GSF, and (2) EKF and CF are close until 2 min. The convergence pattern of these filters also has many similarities. The figures clearly show efficiency of MCF over CF owing to use of optimal weight updates. The error curves for position and velocity are lower for MCF. Moreover, the RMSE criteria (Table: 6-3) and convergence to lower errors shows improvement provided by MCF over other filtering methods. Moreover, the performance of CF is slightly better than EKF. In general the filters based on mixture PDFs (GSF and MCF) show improvement over single approximation of Bayes' *a posteriori* PDF. These error curves are obtained by averaging 50 MC simulations for each filter. This provides a reasonable confidence over these estimation results.

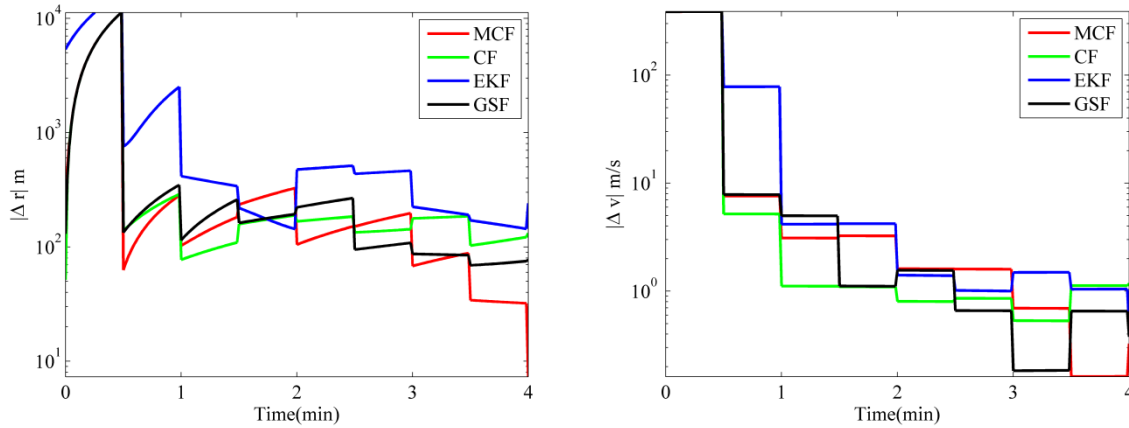


Figure 6-5: Time history of absolute errors (Δ) in ECI coordinates (shown in log scale) for filters with initial conditions of Equation: 6.55 and measurement availability is 0.033 Hz.

Filter	RMSE (Position) (m)			RMSE(Velocity) (m/s)		
	X	Y	Z	\dot{X}	\dot{Y}	\dot{Z}
EKF	2319	2392	2335	80	81	80
CF	1340	1328	1402	79	79	78
MCF	1268	1423	1307	79	79	79
GSF	1298	1390	1316	79	79	79

Table 6-3: RMSE in ECI coordinates for filters with initial conditions of Equation: 6.55 and measurement availability is 0.033 Hz.

On similar lines to Equation: 5.79 one can also define instantaneous RMS error for the filters expressed as [8]:

$$\varepsilon_i(k) = \sqrt{\frac{1}{N} \sum_{j=1}^N [x_{i,j}(k) - \hat{x}_{i,j}(k)]^2} \quad (6.56)$$

where, $x_{i,j}(k)$ is the true state, $\hat{x}_{i,j}(k)$ is the estimated, $i = i^{th}$ component of state, $j = j^{th}$ simulation and N = total number of simulations.

Now one extends the filtering performance for later orbital period i.e., once there are no observations available. Figure: 6-6 to 6-9 depicts time history of IE and instantaneous RMSE (Equation: 6.56) over 3 orbital periods. The measurements (0.033 Hz) are only available for 4 min once the satellite is in viewing position from the radar site. These simulations are obtained from processing 50 MC runs for each filter. The performance of CF and MCF are very close when compared for IE criteria, however, one may observe distinct improvement in RMSE results by MCF over other filtering methods (see Figure: 6-7 to 6-9).

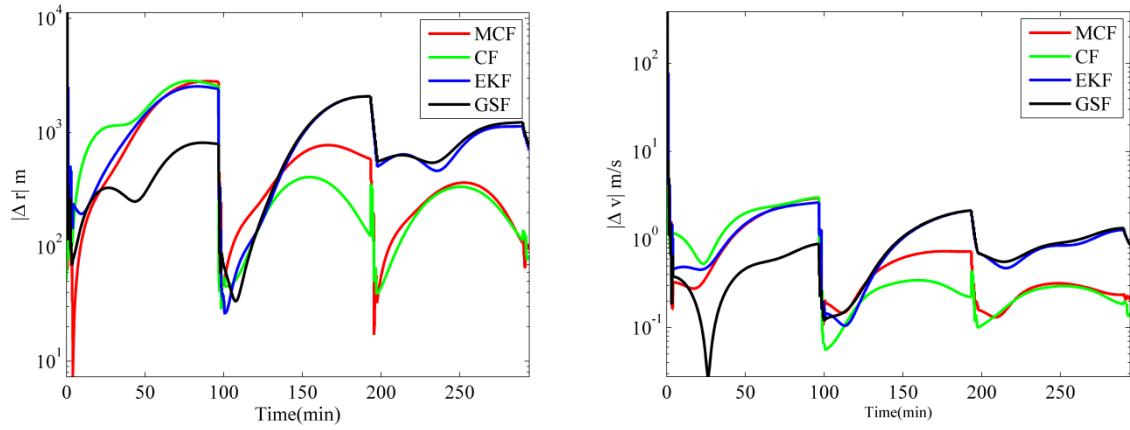


Figure 6-6: Time history of absolute errors (Δ) in ECI coordinates (shown in log scale) for filters with initial conditions of Equation: 6.55 and measurement availability is 0.03 Hz during the satellite is available on horizon only (for 4 min only). This amounts to measurement span once after an orbital period (~ 97 min) during above simulation time i.e., 300 min [5 hr]).

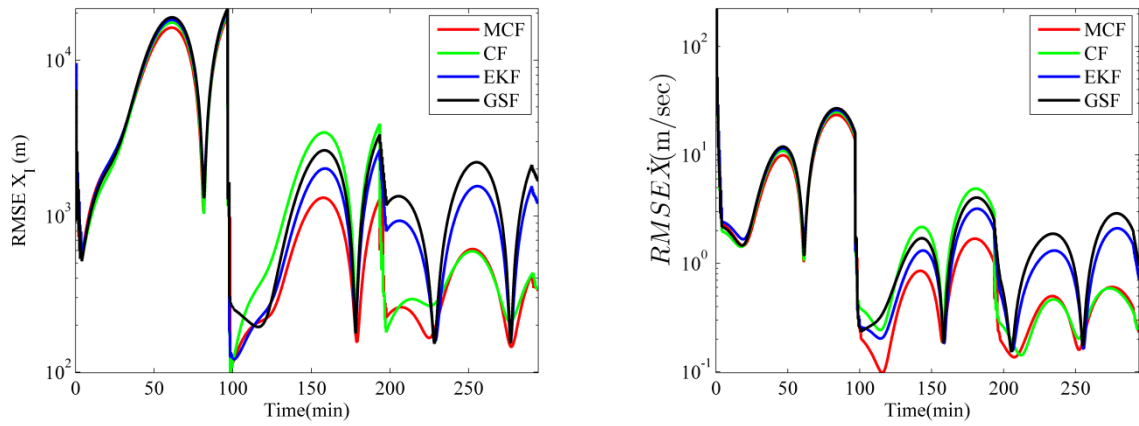


Figure 6-7: Time history of absolute RMSE in ECI X_I (shown in log scale) for filters with initial conditions of Equation: 6.55 and measurement availability is 0.03 Hz during the satellite is available on horizon only (for 4 min only). This amounts to measurement span once after an orbital period (~ 97 min) during above simulation time i.e., 300 min [5 hr]).

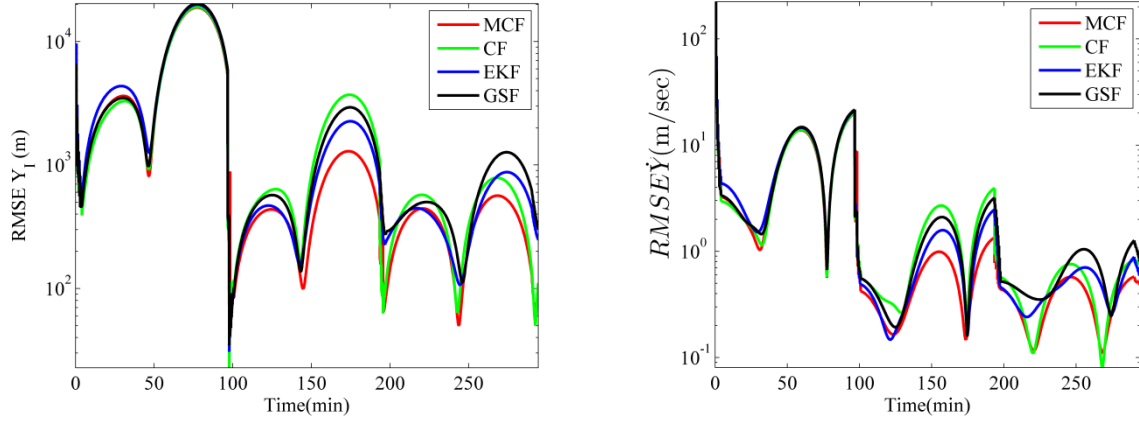


Figure 6-8: Time history of absolute RMSE in ECI Y_I (shown in log scale) for filters with initial conditions of Equation: 6.55 and measurement availability is 0.03 Hz during the satellite is available on horizon only (for 4 min only). This amounts to measurement span once after an orbital period (~ 97 min) during above simulation time i.e., 300 min [5 hr]).

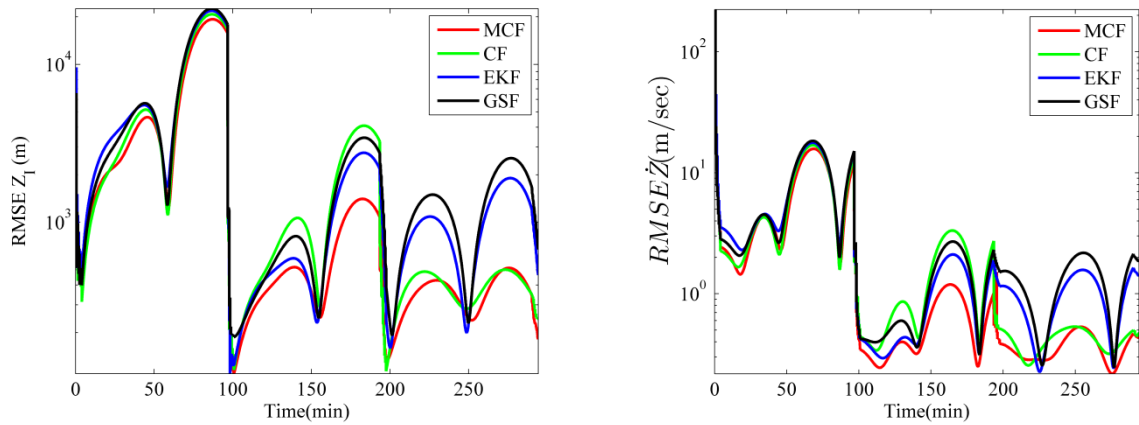


Figure 6-9: Time history of absolute RMSE in ECI Z_I (shown in log scale) for filters with initial conditions of Equation: 6.55 and measurement availability is 0.03 Hz during the satellite is available on horizon only (for 4 min only). This amounts to measurement span once after an orbital period (~ 97 min) during above simulation time i.e., 300 min [5 hr]).

One would now consider a small satellite LEO sun synchronous (inclination = 98 deg) mission consisting of a nano-satellite (weighs less than 10 kg). This satellite is equipped with a GPS receiver i.e., SGR-05P of Surrey Satellite Technology (SSTL). The power requirement for the GPS receiver is 1W at 3.3V and its position and velocity accuracy are 10 m and 0.15 m/s, respectively [122]. Extreme care should be practiced for onboard use of GPS for OD in order to conserve the power and increase satellite's mission lifetime. Therefore, one now extends the use of filters for OD using GPS in a nano-satellite. The position and velocity through GPS device SGR-05P is available after 95 min (~ 1 orbital period). Time history of IE (ECI coordinates) and RMSE for different filters are shown in Figures: 6-10 to 6-13.

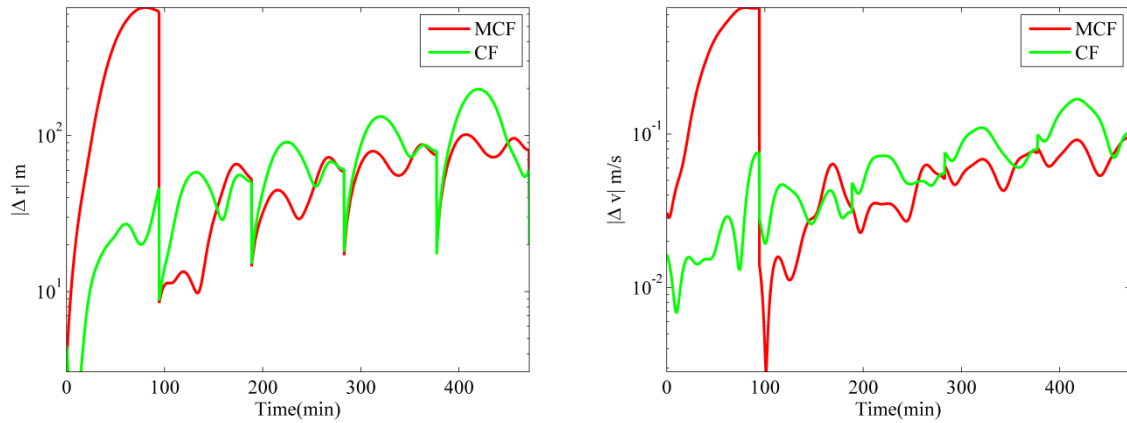


Figure 6-10: Time history of absolute position errors (Δ) in ECI coordinates for filters with initial conditions of Equation: 6.55 and measurement availability is once per orbital period (~ 95 min) using SGR-05P (on board GPS receiver). Simulation time is 500 min (~ 5 orbital periods).

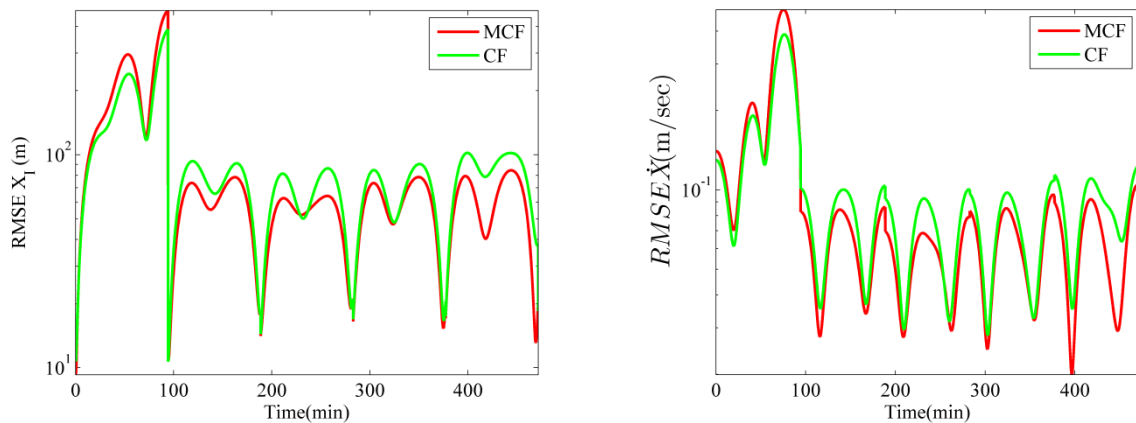


Figure 6-11: Time history of RMSE in ECI coordinates (X-axis) for filters with initial conditions of Equation: 6.55 and measurement availability is once per orbital period (~ 95 min) using SGR-05P (on board GPS receiver). Simulation time is 500 min (~ 5 orbital periods).

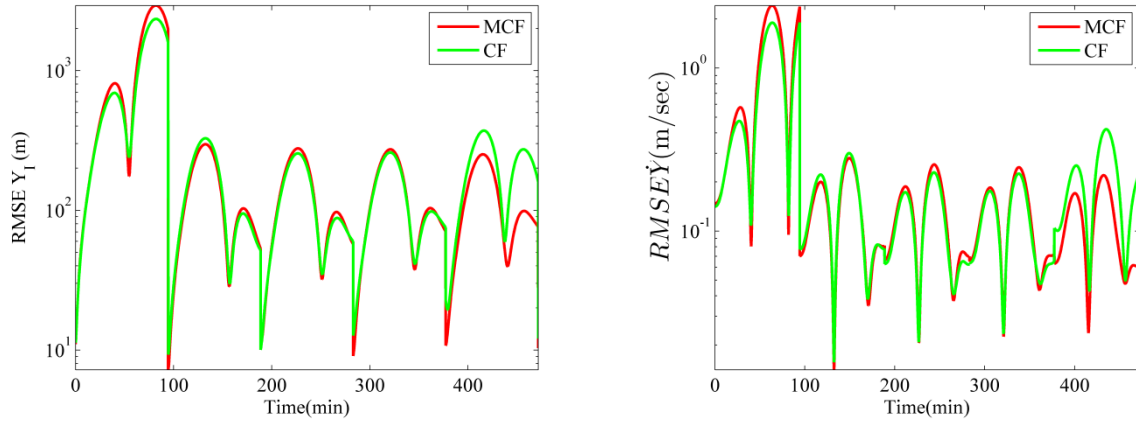


Figure 6-12: Time history of RMSE in ECI coordinates (Y-axis) for filters with initial conditions of Equation: 6.55 and measurement availability is once per orbital period (~ 95 min) using SGR-05P (on board GPS receiver). The simulation time is 500 min (~ 5 orbital periods).

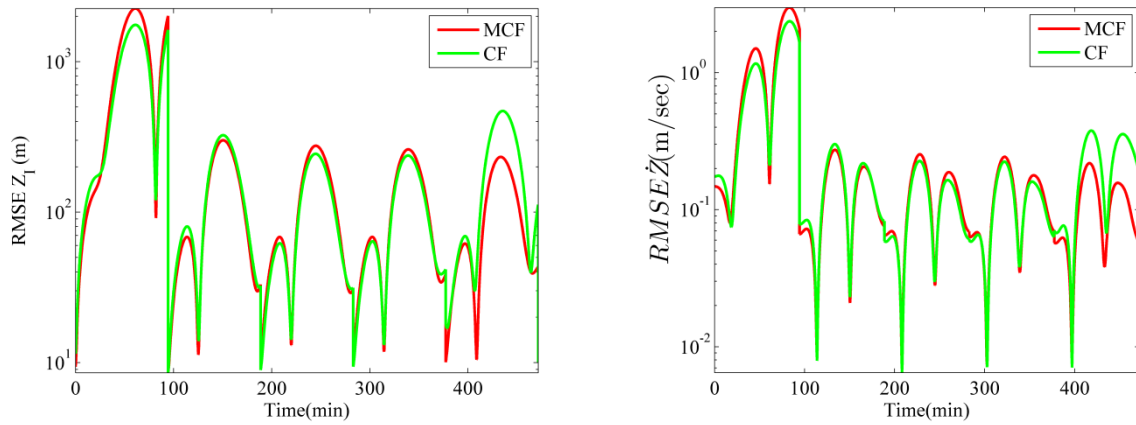


Figure 6-13: Time history of RMSE in ECI coordinates (Z-axis) for filters with initial conditions of Equation: 6.55 and measurement availability is once per orbital period (~ 95 min) using SGR-05P (on board GPS receiver). The simulation time is 500 min (~ 5 orbital periods).

The plots of Figures: 6-10 to 6-13 show that improvement can be achieved using MCF over CF. On careful observation of error plots one can clearly differentiate lower errors produced by MCF in just 5 orbital periods. Although, the errors for MCF are more initially but they quickly (within 5 orbital periods) converge to lower errors. By using MCF one could enhance over all mission life time of a satellite since we are using less measurements which means less power consumed by GPS receiver in addition to better orbit determination. These simulation results have also been produced using 50 MC runs.

6.5.2 Discussion

The results of different filters based on sparse measurements in LEO OD case have been presented. The performance of MCF based on RMSE and averaged errors reveals better estimation accuracy can be achieved especially for extended durations (i.e., multiple orbits). However, its main complexity is due to extensive mathematical derivations such as use of Jacobian, Hessian, Ito calculus, FPKE, and quadratic optimization. The algorithm of CF is comparatively less complex, as it requires calculation of Jacobian, Hessian matrices and use of Ito calculus only to derive multivariate moments. These factors affect the speed of execution, for example based on current Matlab implementation the ratio of time taken by MCF and CF for a particular case is 5:1. This factor does not have much impact for OD based on radar measurements. However, for satellite onboard OD based on GPS, use of MCF would be computationally expensive. On the other hand, *Kalman Filter Framework* provides much simplified implementation due to underlying Gaussian assumption for predictive and Bayes' *a posteriori* PDF. Consequently, the speed of execution is also considerably less, for example the ratio of time taken by MCF and GSF for a particular case is 14:1. Apparently, the ratio appears to be quite significant; however, it does not have much impact on OD based on radar measurements. However, for OD based on GPS measurements the *Kalman Filter Framework* provides significant improvement in computational speed and programming simplicity. Furthermore, better estimation accuracy can be achieved from *Kalman Filter Framework* if one increases the number of measurements. This is mainly due to short term validity of Gaussian approximation for the Bayes' *a posteriori* PDF which may not be optimal under *sparse measurements* environments. Continuous-discrete filtering methodology is more suitable for OD problems due to accessibility of more efficient numerical integration methods such as multistep and extrapolation compared to fixed step RK-4 used by discrete-time filtering. In discrete-time filtering the nonlinear dynamical function (Equation: 6.2) is required to be discretized using smaller time step (order of milliseconds) for high fidelity trajectory generation. This places excessive computational burden on OBC or on ground computers. Whereas, the multistep, extrapolation or variable time step RK methods are more optimal for such requirements due to their better accuracy and speed of execution.

6.6 Lunar Orbital Navigation

The algorithms discussed for *Culver* filter framework would now be implemented for a lunar navigation problem described in reference [1]. In the lunar landing mission the spacecraft will initially be placed in a low altitude circular orbit about the moon. Before the descent phase on to the surface of moon it is extremely important for spacecraft to determine its position and velocity accurately in the lunar orbit [1], in order to avoid landing inaccuracies and damage to spacecraft. This can be accomplished by optically measuring the angles between lunar surface landmarks and the stars at various times, for example in Apollo missions to moon these angular measurements were obtained using sextant [123]. A sextant measures angles between a celestial body like star or a planet and the horizon [124] (It is a traditional equipment used for finding own position during pre-GPS era especially in sea). However, the modern navigational aid for this purpose would be a star tracker.

Firstly one describes the nonlinear orbital system dynamics termed as *true model*. The problem is confined to a planar problem for the sake of simplicity as shown in Figure: 6.16 [1]. The equations of motion in Cartesian coordinate system are expressed as [1]:

$$\ddot{X} = -\frac{\mu_M X}{(X^2 + Y^2)^{3/2}}, \quad \ddot{Y} = -\frac{\mu_M Y}{(X^2 + Y^2)^{3/2}} \quad (6.57)$$

where, $\mathbf{r} = [X, Y]^T$ and $\mathbf{v} = [\dot{X}, \dot{Y}]^T$ are position and velocity vectors respectively, in Cartesian coordinate system of moon and $\mu_M = 4902.799 \text{ km}^3 \cdot \text{s}^{-2}$ (moon gravitational parameter).

The equations describing the discrete time measurements are expressed as:

$$\begin{aligned} \varphi &= \pi + \theta - \alpha \\ &= \pi + \theta - \sin^{-1} \left(\frac{R_L \sin(\theta_l - \theta)}{\sqrt{R_L^2 + r^2 - 2R_L r \cos(\theta_l - \theta)}} \right) \\ \sin(\theta) &= \frac{Y}{r}, \quad \cos(\theta) = \frac{X}{r} \end{aligned} \quad (6.58)$$

where, $r = |\mathbf{r}|$ and other quantities are explained through Figure: 6.14.

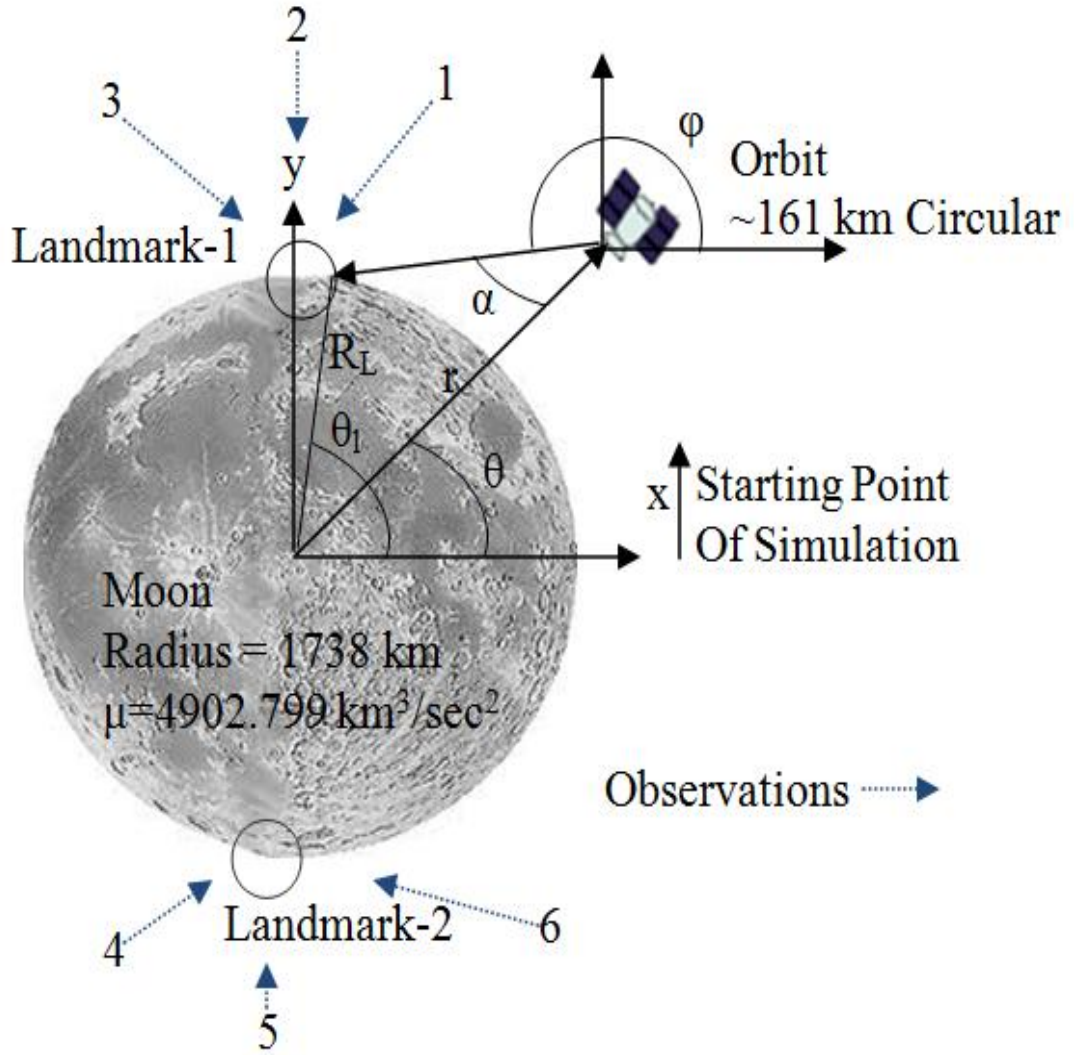


Figure 6-14: Lunar navigation system description

By propagating the states from initial orbit injection data and updating these estimates using discrete optical measurements would form a nonlinear estimation problem. Two landmarks are chosen near polar areas of moon as highlighted in Figure: 6.16 i.e., ± 90 deg. The measurement variance is $R = 0.000016 \text{ rad}^2$ [1]. The state vector to be estimated is:

$$\mathbf{x} = [X, Y, \dot{X}, \dot{Y}]$$

The initial orbit data provided to both the filters is expressed in Equation: 6.59 [1]. CF and MCF (four GCS components) are investigated for eleven orbital periods with observations taken six times per orbit (see Figure: 6-16). The results are as shown in Figure: 6.17-6.18

$$\begin{aligned}
 \hat{\mathbf{x}}_0 &= \begin{bmatrix} 1.8982944 \times 10^3 \text{ km} \\ 0 \text{ km} \\ 0 \frac{\text{km}}{\text{s}} \\ 0.00157734 \times 10^3 \frac{\text{km}}{\text{s}} \end{bmatrix} \\
 \hat{\mathbf{P}}_0 &= \text{diag} \left(\begin{bmatrix} 9.290304 \times 10^{-2} \text{ km}^2, 9.290304 \\ \times 10^{-2} \text{ km}^2, 9.2903040 \times 10^{-8} \frac{\text{km}^2}{\text{s}^2} \end{bmatrix} \right) \\
 \hat{\mathbf{P}}_0^{(3)} &= \hat{\mathbf{P}}_{ijk}^{(3)} = 0
 \end{aligned} \tag{6.59}$$

The Matlab pseudo-random number generator (built in `randn`) is set to default initial state before starting simulations. The initial conditions provided to generate *true trajectory* are slightly different from estimated and are given as:

$$\hat{\mathbf{x}}_0 = \begin{bmatrix} 1.8982944 \times 10^3 \text{ km} \\ 0 \text{ km} \\ 0 \frac{\text{km}}{\text{s}} \\ 0.00157734 \times 10^3 \frac{\text{km}}{\text{s}} \end{bmatrix} + \begin{bmatrix} 0.025 \text{ km} \\ 0.025 \text{ km} \\ 0.0025 \frac{\text{km}}{\text{s}} \\ 0.0025 \frac{\text{km}}{\text{s}} \end{bmatrix}$$

The estimation criteria are taken from Section: 5.7.1 (Equation: 5.78 to 5.80). See IE in Figure: 6-15 and 6-16. These figures depict a significant improvement of MCF over CF. Moreover, the results of RMSE and ME also show better performance of MCF (Tables: 6-4 and 6-5).

Filter	RMSE (km)		RMSE (km/s)	
	ΔX	ΔY	$\Delta \dot{X}$	$\Delta \dot{Y}$
CF	0.7307	0.8863	0.0006	0.0006
MCF	0.4623	0.5288	0.0004	0.0004

Table 6-4: RMSE for a lunar navigation problem

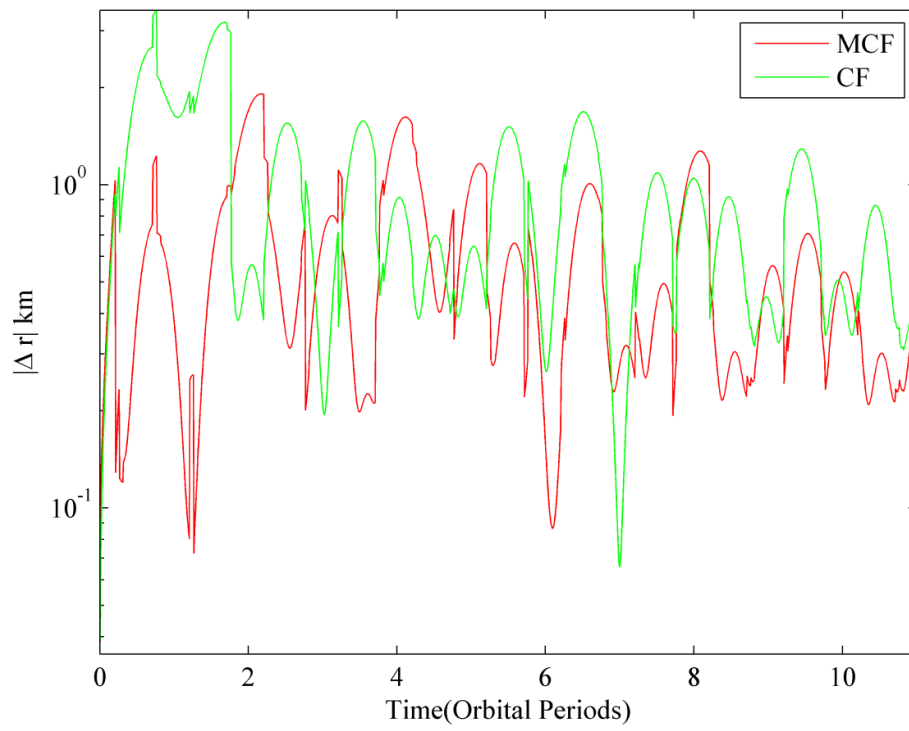


Figure 6-15: Time history of absolute position errors (Δ) in Cartesian positions for Culver framework under sparse measurements.

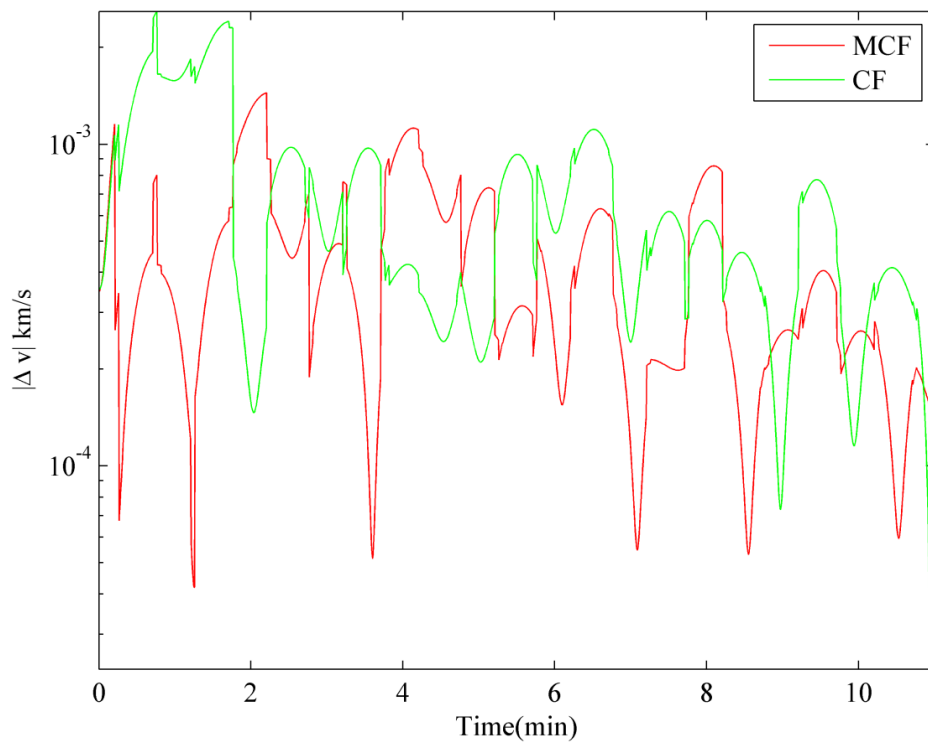


Figure 6-16: Time history of absolute velocity errors (Δ) in Cartesian velocities for Culver framework under sparse measurements.

Filter	ME (km)		ME (km/s)	
	ΔX	ΔY	$\Delta \dot{X}$	$\Delta \dot{Y}$
CF	-0.0658	0.6013	-0.0001	0.0
MCF	-0.1912	0.3474	-0.0001	0.0

Table 6-5: ME for a lunar navigation problem

6.7 Summary

In this chapter a detailed description of filters based on KF and CF have been described. A new filter based on GCSMM namely MCF has been presented. The algorithm has been compared with EKF, GSF, and CF for nonlinear Earth satellite OD through ground based radar system, GPS and with CF for a lunar orbital navigation problem. The results show some improvements in RMSE and ME by MCF. Therefore, MCF and CF can be considered as better choice for replacement of Gaussian PDF based nonlinear filters especially under sparse measurements and highly uncertain initial conditions.

7 Conclusion and Future Work

7.1 Introduction

This chapter provides a brief overview of the work presented in the thesis emphasizing notable results. The main theme of this research pertains to Bayesian estimation of nonlinear dynamical systems using a mixture of orthogonal expansions along with its applications for sequential orbit determination of space objects around Earth. In addition, the non-Bayesian approach to estimation i.e., least squares has also been used with a view to carry out analysis of fidelities of LEO absolute and relative motion models [37] and long term parameter estimation of Epicyclic orbits [38].

7.2 Concluding Summary

The analytical description of the dynamics of satellites in Earth orbits has been a focus of intense research since 50 years. Analytical description provides a clearer physics of the underlying motion. However, due to neglect of unmodelled dynamics and linearization these models only approximate truth. In Chapter: 3 nonlinear least squares or GLDC scheme has been adapted to analyze fidelities of analytical models by estimating initial conditions of absolute and relative motion models. The results have shown that the estimated initial conditions significantly improve the analytic orbit propagation accuracy for longer time i.e., weeks. Moreover, high fidelity models like Epicycle [2] or J_2 modified HCW [20] have less free propagation errors even after batch least squares fitting span, compared to the unperturbed Kepler's two body problem [13] or simple HCW equations [18]. These initial conditions can be used to incorporate conservative (e.g., zonal or tesseral harmonics of geopotential) or non-conservative (e.g., atmospheric drag) perturbative effects in satellite feedback control systems or simply high precision trajectory propagations.

In Chapter: 4 a new parameter estimator for Epicyclic orbits [38] has been derived which exploits linear secular perturbative effects in Epicycle orbital coordinates of argument of latitude and right ascension of the ascending node. The accuracy achieved using EPF can easily be extended for higher order zonal perturbative terms. The estimation results show improved epicycle coordinates compared to the nonlinear numerical trajectory. It was found out that by keeping drift in the mean errors as 10% of the maximum error in a particular position coordinate, repeated estimation of the epicycle parameters would be needed after *twelve days*.

In Chapter: 5 SMC methods based on weighted point mass approximation of Bayes' *a posteriori* PDF have been used to extend more optimal *parameteric bootstrap* PFs using GCS, GCSMM and CF Hybrids. These filters employ full GCS (complete higher order moments [order ≤ 5] including cross

moments) and marginal GCS (axial moments for $[\text{order} \geq 3]$) using Gaussian Copula [112] in mixture configuration. To author's knowledge this is first attempt of a unified presentation of GCS and GCSMM for discrete time SMC estimation. The algorithms have been simulated on simple nonlinear pendulum and orbit determination of satellite in LEO using radar measurements. The results show improved Instantaneous Error (IE) and Root-Mean-Square-Error (RMSE) for both the nonlinear problems when compared with generic PF, GPF, GSPF and EKF. Nevertheless, GCS based algorithms are more complex and slightly time inefficient (see Table: 5.4).

In Chapter: 6 *continuous-discrete* filtering of nonlinear dynamical system has been reviewed. A new filtering algorithm based on GCSMM has been developed called as MCF. The filter specializes use FPKE error feedback methodology [120] for time update of GCSMM weights. A framework for filters i.e., *Kalman and Culver Filters* is formed for qualitative and quantitative analysis. MCF has shown better performance in terms of IE and RMSE when compared with other filters.

In view of results presented in Chapter: 5 and 6 it can be ascertained that nonlinear Bayesian filtering based on mixture of orthogonal expansion is more optimal than single expansions. In particular GCS of lower order could be suboptimal representation of true non-Gaussian PDFs (see Figure: 1-3 and 5-6). Use of such PDFs for nonlinear estimation could possibly lead to divergence and sub optimal uncertainty quantification such as non positive definite covariance matrices. Increase in the order of GCS could probably lead to better performance. However, it would be at the cost of tremendous complexity and extensive computation. Therefore use of mixture of lower order GCS for nonlinear estimation is deemed more suitable for performance enhancement in nonlinear estimation applications.

7.3 Research Achievements

The research achievements can be summarized as under:

- Development of new GCSPF, GCSMPF and hybrids for nonlinear Bayesian discrete-time state estimation based on MC simulation approach [41],[45],[42]. The filters have shown improvement over other filtering methods such as EKF and generic Particle Filter (PF) under *sparse measurements* availability.
- Development of new filter namely MCF, based on third ordered GCSMM approximation of the Bayes' *a posteriori* PDF. More particularly it utilizes optimal FPKE error feedback, hybrid analytical and numerical (i.e., quadratic optimization) to compute weights associated with each component of GCSMM.
- The application of new nonlinear Bayesian filters are simulated for simple pendulum, LEO satellite OD and navigation of lunar orbiter under *sparse measurements* and compared with other state of the art nonlinear filters such as EKF, GSF, GPF, GSPF, CF and PF (SIS-R). This provides a unified view on use of GCS and GCSMM for nonlinear

state estimation based on Taylor series and MC simulations.

- Analysis on fidelities of linearized LEO absolute and relative motion models namely Kepler's equation, Epicycle model around oblate Earth, HCW and SS equations using GLDC scheme [46],[39],[40]. The selection of appropriate IC or parameters of analytic models is vital to minimize the process noise and obtain more accurate orbital trajectories for such comparisons.
- Development of a new algorithm based on linear least squares for parameter estimation of Epicyclic orbit namely EPF. The method exploits naturally occurring linear secular increase in Epicyclic coordinates of argument of latitude and RAAN. The estimated parameters enable minimization of the process noise and long term high fidelity orbital trajectory generation at all inclinations for LEO [38].

7.4 Extensions and Future Work

Nonlinear estimation is a challenging and growing field. With increase in computational resources faster algorithms can be developed using numerical solutions of FPKE and Bayes' formula. However, in near future they can only be used for ground based applications. Based on work in this thesis possible direction of future research are:

- (1) Further to the work presented in Chapter: 3 one may develop more realistic models of the process noise using differential or difference equations. For example consider use of Matlab system identification toolbox, linear state space modelling of dynamical systems based on least squares i.e., the function `n4sid`. In general these models could provide performance enhancement of batch or sequential filters for any nonlinear estimation requirement.
- (2) Enhancement of fidelity of MCF by using mixture of higher order GCS expansions. However, it is suggested that these expansions should be attempted for low dimensional systems i.e., (dimension ≤ 2). As analytical solutions for higher dimensional system could be quite extensive.
- (3) It is envisaged that extended propagation of satellite nonlinear dynamics over several orbital periods / days under *sparse measurements*, using MC based algorithms such as GCSPF and GCSMPF of lower order GCS i.e., (order ≤ 3) would be insufficient to capture non-Gaussianity of state distribution. Therefore, higher order GCS (order > 3) copula based random vector generators be developed for long term prediction and approximation of state distributions.
- (4) GCS random number generator based on AR method (see Chapter: 5), in comparison to Gaussian random generators available in Matlab is suboptimal with regards to speed of execution and percentage (approximately 25-30%) rejection of unusable random vectors. This poses a serious issue in real time filtering for higher dimensional (dimension ≤ 2) problems.

Therefore, it is recommended that more efficient GCS random number generators be developed based on analytical methods such as inverse transform method [67] instead of MC simulation based AR.

- (5) The more optimal GCSMM approximation of satellite state predictive PDF is useful for long term (i.e., months) predictions without any measurements. Thus, providing better state uncertainty quantification of such dynamical systems. With growing number of satellites in LEO, there is a requirement to carry out space object conjunction analysis. It is suggested that methodology of Section: 6.4.1, be used to obtain predictive PDF for computation of probability of likely collisions between satellites or space debris.
- (6) The nonlinear filters based on GCS and GCSMM are computationally more expensive than Gaussian based filters such as EKF or GSF. More specifically, methodology of Section: 6.4.1 is currently not optimized for use on OBC due to computing resource limitations on satellites. Therefore, as future works the algorithm be optimized in C++ for OBC requirements and tested on space qualified hardware such as A712 OBC for STRanD-1 [125].

References

- [1] Culver, C. O., "Optimal Estimation for Nonlinear Stochastic Systems," PhD Dissertation Massachusetts Institute of Technology, Mar 13, 1969.
- [2] Hashida, Y., Palmer, P., "Epicyclic Motion of Satellites About an Oblate Planet," *Journal of Guidance Control and Dynamics*, Vol 24, No.3, 2001, pp 586-596.
- [3] Maybeck, P., "Stochastic Models, Estimation, and Control Vol. 1," Academic Press London 1979.
- [4] Scheinerman, E.R., "Invitation to Dynamical Systems," Department of Mathematical Sciences, John Hopkins University, Page. 1, Jun 2000.
- [5] Shalom, Y.B., Li, X.R., Kirubarajan, T., "Estimation with Application to Tracking and Navigation," Wiley-Interscience Publication 2001.
- [6] Jazwinski, A. H., "Stochastic Process and Filtering Theory," Published by New York Academic, 1970.
- [7] <http://www.answers.com/topic/low-earth-orbit> [Viewed on: 26 Jul 2011].
- [8] Lee, D.J., "Nonlinear Bayesian Filtering with Applications to Estimation and Navigation," PhD Dissertation, Texas AandM University May 2005.
- [9] Wertz, J.R., Larson, W.J., "Space Mission Analysis and Design", Space Technology Library, published jointly by Microcosm Press, 1999.
- [10] Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," *Transaction of the AMSE Journal of Basic Engineering*, 82 (D-Series): pages: 35-45 1960.
- [11] Horwood, J. T., Aragon, N. D., Poore, A. B., "Edgeworth Filters for Space Surveillance Tracking," Advanced Maui Optical Space Surveillance Technologies Conference 2010.
- [12] Vallado, D. A., "Fundamentals of Astrodynamics and Applications (Second Edition)," The Space Technology Library, Published by Microcosm Press and Kluwer Academic Publisher 2004.
- [13] Montenbruck, O., Gill, E., "Satellite Orbits Models, Methods and Applications," Published by Springer, 2005.
- [14] Hashida, Y., "Analytical Solutions for Autonomous Determination of Near Circular Orbits," PhD Dissertation Surrey Space Centre April 2003.
- [15] Kozai, Y., "The Motion of a Close Earth Satellite," *Astronomical Journal*, Vol. 64, No.1274, 1959, pp. 367-377.
- [16] Brouwer, D., "Solution of the Problem of Artificial Satellite Theory Without Drag," *Astronomical Journal*, Vol. 64, No. 1274, 1959, pp. 378-397.
- [17] Zink, M., K. Gerhard., Fiedler, H., et.al., "TanDEM-X – The First Bistatic SAR Formation in Space," Geoscience and Remote Sensing Symposium 2007, IGARSS 2007, IEEE International.
- [18] Hill, G. W., "Researches in the Lunar Theory," *American Journal of Mathematics*, Vol. 1, 1878,

pp. 5–26.

- [19] Clohessy, W. And Wiltshire, R., “Terminal Guidance System for Satellite Rendezvous,” *Journal of the Astronautical Sciences*, Vol. 27, No. 9, 1960, pp. 653–678.
- [20] Schweighart, A. S., Sedwick, R. J., “High Fidelity Linearized J_2 Model for Formation Flight,” *Journal of Guidance Control and Dynamics*, Vol 25, No.6, pages. 1073-1080, 2002.
- [21] Sarkka, S., “Recursive Bayesian Inference on Stochastic Differential Equations,” PhD Dissertation 2006, Helsinki University of Technology Laboratory of Computational Engineering.
- [22] Merwe, R.V.D., “Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models,” PhD Dissertation OGI School of Science and Engineering at Oregon Health and Science University, Apr 2004.
- [23] Maybeck, P., “Stochastic Models Estimation and Control Vol. 2,” Published by Academic Press Inc. 1979.
- [24] Risken, H., “The Fokker-Planck Equation,” Second Edition, Springer 1996.
- [25] Daum, F., “Nonlinear Filters: Beyond the Kalman Filter,” IEEE Aerospace and Electronics Magazine Vol. 20, No. 8, Aug 2005.
- [26] Stuart, A., Ord, J. K., “Kendall’s Advanced Theory of Statistics,” Vol. 1 Distribution Theory, 6th Edition, Copublished in Americas by Halsted Press an Imprint of John Wiley and Sons Inc, 1994.
- [27] Blinnikov, S., Moessner, R., “Expansions for Nearly Gaussian Distribution,” *Astronomy and Astrophysics Supplement Series*. 130, pages. 193-205, May 1998.
- [28] McCullagh, P., “Tensor Methods in Statistics,” Publisher Chapman and Hall, 1987.
- [29] Charlier, C,V,L., “Über Dir Darstellend Willkürlicher Funktionen,” *Ark.Mat Astr.och Fys.*2 No.20 1-35, 1906.
- [30] Bucy, R,S., Hecht, C., Captain Senne, K,D., “An Engineer’s Guide to Building Nonlinear Filters,” USAF Final Project Report No. 7904-00-17, May 1972.
- [31] <http://mathworld.wolfram.com/HermitePolynomial.html> [Viewed on: 19 Jul 2011].
- [32] Edgeworth, F. Y., “The Law of Error,” *Cambridge Philos. Soc.* 20, 36-66 and 113-141, 1905.
- [33] Challa, S., Shalom, Y. B., Krishnamurty, V., “Nonlinear Filtering via Generalized Edgeworth Series and Gauss-Hermite Quadrature,” *IEEE Transaction on Signal Processing*, Vol.48, No.6, pages. 1816-1820, Jun 2000.
- [34] Van Hulle, M, M., “Edgeworth-Expanded Gaussian Mixture Density Modeling,” *Journal of Neural Computation*, Vol. 17, No. 8, Pages. 1706-1714, Aug 2005.
- [35] Gilani, S.A.A., Palmer, P., “Nonlinear Bayesian Estimation Based on Mixture of Gram Charlier Series”, Proceedings of IEEE Aerospace Conference, Mar 2012.
- [36] Muscolino, G., Ricciardi, G., Vasta, M., “Stationary and Non-Stationary Probability Density Function for Non-Linear Oscillators,” *International Journal of Nonlinear Mechanics* Vol.32, No. 6, pp. 1051-1064, 1997, Elsevier Science Ltd.
- [37] Gilani, S.A.A., Palmer, P.L., “Analysis of Fidelities of Linearized Orbital Models using Least

- Squares,” Annual IEEE Aerospace Conference at Big Sky, USA, Mar 2011.
- [38] Palmer, P.L., Gilani, S.A.A., “Epicycle Parameter Filter for Long Term Orbital Parameter Estimation,” 25th Annual USU/AIAA Small Satellite Conference, Utah, USA, Aug 2011.
 - [39] Gauss, C. F., “Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium,” Published 1809 by Sumtibus F. Perthes et I.H.Besser in Hamburgi.
 - [40] Yan, H., Alfriend, K.T., Vadali, S.R., Segupta, P., “Optimal Design of Satellite Formation Relative Motion using Least Squares,” *Journal of Guidance Control and Dynamics*, Vol. 32, No. 2, pages. 599-604, Mar-Apr 2009.
 - [41] Gordon, N., Salmond, D., Ewing, C., “Bayesian State Estimation for Tracking and Guidance using the Bootstrap Filter,” *Journal of Guidance Control and Dynamics*, Vol. 18, No. 6, pages. 1434-1443, Nov-Dec 1995.
 - [42] Kotecha, J. H., Djuric, P. M., “Gaussian Particle Filtering,” *IEEE Transactions on Signal Processing*, Vol. 51, No. 10, pages. 2592-2601, Oct 2003.
 - [43] Kotecha, J. H., Djuric, P. M., “Gaussian Sum Particle Filtering for Dynamic State Space Models,” Proceedings of ICASSP-2001, Salt Lake City, Utah, May 2001.
 - [44] Alspach, D. L., Sorenson, H. W., “Nonlinear Bayesian Estimation using Gaussian Sum Approximation,” *IEEE Transaction on Automatic Control*, Vol. AC-17, No.4, pages. 439-448, Aug 1972.
 - [45] Doucet, A., Godsill, S., and Andrieu, C., “On Sequential Monte Carlo Sampling Methods for Bayesian Filtering,” *Statistical Computing* Vol. 10, No.3, pp. 197-208, 2000.
 - [46] Gauss, C.F., Translated by Stewart, G.W., 1995, “Theory of the Combination of Observations Least Subject to Errors: Part One, Part Two, (Supplement)”, Philadelphia: Society for Industrial and Applied Mathematics.
 - [47] Ristic, B., Arulampalam, S., Grodon, Neil., “Beyond the Kalman Filter: Particle Filters for Tracking Applications,” Published by Artech House, 2004.
 - [48] Julier, S. J., and Uhlmann, J. K., “A New Extension for Kalman Filter to Nonlinear Systems,” In The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Orlando, FL, USA, 1997. SPIE. Multi Sensor Fusion, Tracking and Resource Management II.
 - [49] Terejanu, G., Singla, P., Singh, T., and Scott, P.D., “A Novel Gaussian Sum Filter Method for Accurate Solution to Nonlinear Filtering Problem”, IEEE 11th International Conference on Information Fusion, 2008.
 - [50] De Freitas, J.F.G., “Bayesian Methods for Neural Networks,” PhD Thesis Cambridge University Engineering Department, 1999.
 - [51] Kastella, K., Kreucher, C., Pagels, M. A., “Nonlinear Filtering for Ground Target Applications,” In Signal Data Processing of Small Targets 2000, Proceedings of SPIE Vol: 4048 (2000).
 - [52] Beal. M.J., “Variational Algorithms for Approximate Bayesian Inference”, PhD Dissertation,

University of London, May 2003.

- [53] Grewal, M.S., Andrews, A.P., “Applications of Kalman Filtering in Aerospace 1960 to Present”, *IEEE Control Systems Magazine*, Jun 2010.
- [54] Ito, K., and Xiong, K., “Gaussian Filters for Nonlinear Filtering Problems”, *IEEE Transaction on Automatic Control*, Vol. 45, No. 5, pages. 910-927, May 2000.
- [55] Norgaard, M., Poulsen, N.K., Ravn, O., “Advances in Derivative-Free Estimation of Nonlinear Systems”, Report No. IMM-REP-1998-15, Technical University of Denmark, Apr 2000.
- [56] Stirling, J., “Methodus Differentialis, Sive Tractatus De Summation Et Interpolation Serierum Infinitarum”, 1730.
- [57] Merwe, R.V.D., Wan, E.A., “The Square-root Unscented Kalman Filter for State and Parameter Estimation”, IEEE International Conference on Acoustic Speech and Signal Processing, 2001.
- [58] Merwe, R.V.D., Wan, E.A., “Efficient Derivative-Free Kalman Filters for Online Learning”, European Symposium on Artificial Neural Networks, ISBN 2-930307-01-3, pp. 205-210, Burges (Belgium), Apr 2001.
- [59] Lee, D.J., Alfriend, K.T., “Sigma Point Filtering for Sequential Satellite Orbit Estimation and Prediction”, *Journal of Spacecraft and Rockets*, Vol. 44, No. 2, pages. 388-398, Mar-Apr 2007.
- [60] Anderson, B, D, O., Moore, J, B., “Optimal Filtering,” Prentice-Hall Information and System Science Series, 1979.
- [61] Ali-Loytty, S., “Gaussian Mixture Filters in Hybrid Positioning”, PhD Dissertation Aug 2009, Tampere University of Technology.
- [62] Horwood, J., Poore, A., “Adaptive Gaussian Sum Filters for Space Surveillance”, IEEE Transactions on Automatic Control, 2010.
- [63] Terejanu, G., Singla, P., Singh, T., and Scott, P.D., “Adaptive Gaussian Sum Filters for Nonlinear Bayesian Estimation”, IEEE Transactions on Automatic Control, Vol.56, Issue.9, pages: 2151-2156, Sep 2011.
- [64] Liu, J.S., Chen, R., “Sequential Monte Carlo Methods for Dynamic Systems”, *Journal of American Statistical Association*, 93 (1998), 1032-1044.
- [65] Gang, W., Xiao-Jun, D., “Particle Filtering and its Application in Satellite Orbit Determination”, IEEE Congress on Image and Signal Processing, 2008.
- [66] Musso, C., Oudjane, N and LeGland, F., “Improving Regularised Particle Filters”, in *Sequential Monte Carlo Methods in Practice*, by Doucet, A., de Freitas, J. F. G., and Gordon, N. J., New York, Springer-Verlag, 2001.
- [67] Martinez, W.L., Martinez, A.R., “Computational Statistics Handbook with MATLAB”, Published by Chapman and Hall/CRC, London, 2002.
- [68] Arulampalam, M.S., Maskell, S., Gordon, N., Clapp, T., “A Tutorial for Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking,” *IEEE Transaction on Signal Processing* Vol:50, No:2, pages. 174-188, Feb 2002.

- [69] Daum, F., Huang, J., “Nonlinear Filtering with Quasi-Monte Carlo Methods”, Signal and Data Processing of Small Targets 2003, Edited by Oliver E. Drummond, Proceedings of SPIE Vol. 5204.
- [70] Murphy, K., and Russell, S., “Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks,” in Sequential Monte Carlo Methods in Practice, edited by Doucet, A., de Freitas, N., and Gordon, N. J., Springer, New York, 2001.
- [71] Nordlund, P.J., “Sequential Monte Carlo Filters and Integrated Navigation”, Thesis Department of Electrical Engineering, Linkopings University, Linkopings, Sweden, 2002.
- [72] Pitt, M., and Shephard, N., “Filtering via Simulation: Auxiliary Particle Filters,” *Journal of American Statistical Association*, Vol. 94, No. 446, pp. 590–599, 1999.
- [73] Cameron, R.H., and Martin, W.T., “The Orthogonal Development of Nonlinear Functionals in Series of Fourier-Hermite Function”, *Ann. of Math.* 48, 385-392(1947).
- [74] Kuznetsov, P.I., Stratonovich, R.L., and Tikhonov, V.I., “Quasi-Moment Functions in the Theory of Random Processes”, *Theory Probab. Appl.* 5(1), 80-97(1960).
- [75] Srinivasan, K., “State Estimation by Orthogonal Expansion of Probability Distribution”, *IEEE Transaction on Automatic Control* AC-15 (1), 3-10(1970).
- [76] Kizner, W., “Optimal Nonlinear Estimation Based on Orthogonal Expansion”, Technical Report 32-1366, JPL California Institute of Technology, NASA, 1969.
- [77] Abramowitz, M., Stegun, I.A., “Handbook of Mathematical Functions With Formulas, Graph and Mathematical Tables”, National Bureau of Standards and Applied Mathematics, 1964.
- [78] Yang, W.Y., Cao, W., Chung, T.S., and Morris, J., “Applied Numerical Methods using Matlab”, John Wiley and Sons, Inc., Publication, 2005.
- [79] Griebel, M., Schweitzer, M.A., “Meshfree Methods for Partial Differential Equations”, Published by Springer, 2002.
- [80] Daum, F.E., “Exact Finite Dimensional Nonlinear Filter”, *IEEE Transaction on Automatic Control*, Vol. AC-31, No. 7, Jul 1986.
- [81] Tapley, B.D., Schutz, B.D., Born, G.H., “Statistical Orbit Determination”, Published by Elsevier Academic Press, 2004.
- [82] Legendre, A.M., “Nouvelles méthodes pour la détermination des orbites des comètes (1806; “New Methods for the Determination of Comet Orbits”)", 1806.
- [83] Donahue, W. H., “New Astronomy,” Cambridge University Press, 1992 translation of Kepler, J., “*Astronomia Nova...*”, Heidelberg, 1609.
- [84] Alton, E.J., Duncan, A.M., Field, J.V., “The Harmony of the World,” *Memoirs of American Philosophical Society*, Vol.209, Philadelphia, 1997, translation of Kepler, J., “*Harmonices Mundi Libri V.Linz*, 1619.
- [85] Battin, R., “An Introduction to the Mathematics and Methods of Astrodynamics,” AIAA Education Series, Jun 1987.

- [86] Braukhane, A., Arza, M., Bacher, M., Calaprice, M., Fiedler, H., Koehne, V., McGuire, H.R., Rivera, J.J., "FormSat, a Scalable Formation Flying Communication Satellite System", IEEE Aerospace Conference, 2010.
- [87] <http://trs-new.jpl.nasa.gov/dspace/bitstream/2014/9712/1/02-1880.pdf> [Viewed on 27 Jul 2011].
- [88] Vaddi, S. S., "Modelling and Control of Satellite Formations," PhD Dissertation Texas AandM University May 2003.
- [89] Lawden, D. F., *Optimal Trajectories for Space Navigation*, Butterworths, London, 1963, pp. 79-86.
- [90] Tschauner, J., Hempel, P., "Rendezvous zu ein min Elliptischer Bahn Umlaufenden Ziel," *Astronautica Acta*, Vol. 11, No. 2, 1965, pp. 104-109.
- [91] Halsall, M., Palmer, P., "Modelling Natural Formations of LEO Satellites," *Celestial Mech Dyn Astr* (2007) Springer Science.
- [92] http://www.centennialofflight.gov/essay/Dictionary/GEO_ORBIT/DI146.htm [Viewed on 15 Aug 2011].
- [93] King-Hele, D.G., "The Effect of Earth's Oblateness on the Orbit of a near Satellite," *Proceedings of the Royal Society of London, Series A: Mathematical and Physical Sciences*, Vol.247, 1958, pp.49-72.
- [94] Palmer, P., "Reachability and Optimal Phasing for Reconfiguration in Near Circular Orbit Formations" AIAA.
- [95] Palmer, P., "Optimal Relocation of Satellites Flying in Near-Circular-Orbit Formulations," *Journal of Guidance Control and Dynamics*, Vol 29, No.3, 2006, pp 519-526.
- [96] Irvin, J. D., Cobb, G, Richard., "Fuel-Optimal Maneuvers for Constrained Relative Satellite Orbits," Air Force Institute of Technology, Dayton, Ohio.
- [97] Han. H., Sparks. A., "Geometry and Control of Satellite Formations," *Proceedings of American Control Conference*, Chicago, Illinois, Jun 2000.
- [98] Lim. H. C etal., "Tracking Control Design Using Sliding Mode Techniques for Satellite Formation Flying," *Journal of Astronomical Space Sciences*, 20(4), 365-374, 2003.
- [99] Kay, S.M., "Intuitive Probability and Random Processes using Matlab", Published by Springer, 2006.
- [100] Thomas, G.B., Finney., R.L., "Calculus of Analytical Geometry," 9th Edition, Addison-Wesley Publishing Company, Jun 1998.
- [101] <http://celestrak.com/NORAD/elements/> [Viewed on 25 Oct 2010].
- [102] Surrey Satellite Technology Limited (SSTL), EADS Astrium, <http://www.sstl.co.uk/> [Viewed on 23 May 2011].
- [103] Haug. A.J., "A Tutorial on Bayesian Estimation and Tracking Technique Applicable to Nonlinear and Non-Gaussian Processes", MITRE Technical Report No. MTR 05W0000004, Jan 2005.

- [104] <http://mathworld.wolfram.com/DeltaFunction.html> [Viewed on 2 Sep 2011].
- [105] De Frietas. J.G.G., “Bayesian Methods for Neural Networks”, PhD Dissertation, Engineering Department, University of Cambridge.
- [106] Spiring. F., “The Refined Positive Definite and Unimodal Regions for the Gram-Charlier and Edgeworth Series Expansion”, *Journal of Advances in Decision Sciences*, Vol. 2011(2011), Article Id. 463097.
- [107] Dempster. A.P., Liard. N.M., and Rubin. D.B., “Maximum Likelihood for Incomplete Data via the EM Algorithm”, *Journal of Royal Statistical Society B*, 39, 1-38.
- [108] Redner. R.A., and Walker. H.F., “Mixture Densities, Maximum Likelihood and the EM algorithm”, *SIAM Rev.*, 26, 195-239.
- [109] Nabney. I.T., “Netlab Algorithms for Pattern Recognition”, Published by Springer-Verlag, 2004.
- [110] <http://www.cs.duke.edu/courses/spring04/cps196.1/handouts/EM/tomasiEM.pdf> [Viewed on: 24 Sep 2011].
- [111] Greiner. W., “Classical Mechanics: Systems of Particles and Hamiltonian Dynamics”, Springer-Verlag, 2003.
- [112] Nelson. R.B., “An Introduction to Copulas”, Published by Springer, 2006.
- [113] <http://www.iro.umontreal.ca/~eckdoug/vibe/Harmonic/PendulumDamped.html> [Viewed on 6 Sep 2011].
- [114] Higham. D.J., “An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations”, *SIAM Review* Vol. 43. No. 3, pp. 525-546.
- [115] Smith. O.E., Adelfang. S.I., and Tubbs. J.D., “A bivariate Gamma Probability Distribution with Application to Gust Model”, NASA Technical Memorandum, TM-82483.
- [116] Rawat. T.K., and Parthasarathy., “Satellite Tracking using a Second-Order Stochastic Nonlinear Filter”, *HAIT Journal of Science and Engineering B*, 2007 Holon Institute of Technology.
- [117] Grewal. M.S., Weill. L.R., Andrews. A.P., “Global Positioning System Inertial Navigation and integration”, Wiley-Interscience Publication, 2007.
- [118] <http://yachtpals.com/how-to/celestial-navigation> [Viewed on: 12 Sep 2011]
- [119] Terejanu. G., Singla. P., et.al., “Uncertainty Propagation for Nonlinear Dynamic Systems Using Gaussian Mixture Model”, *Journal of Guidance Control Dynamics*, Vol.31, No.6, pages. 1622-1633, Nov-Dec 2008.
- [120] Terejanu. G., “Towards a Decision Centric Framework for Uncertainty Propagation and Data Assimilation”, PhD Dissertation 2010, State University of New York at Buffalo.
- [121] Kerr, T.H., “Streamlining Measurement Iteration for EKF Target Tracking”, *IEEE Transaction on Aerospace and Electronic Systems*, Vol.27, No.2, Mar 1991.
- [122] Ebinuma, T., Unwin, E., Underwood, M and Imre., C, “A miniaturized GPS receiver for space

application”, IFAC Automatic control in Aerospace, 2004.

- [123] http://www.ion.org/museum/item_view.cfm?cid=6andscid=5andiid=293 [Viewed on: 16 Oct 2011].
- [124] Bauer. B., “A Handbook of Sextant”, Published by International Marine, an imprint of McGraw-Hill Inc, 1995.
- [125] Kenyon, S., Bridges, C., “Use of smart phone as a central avionics of a nanosatellite”, paper appeared in 62 International Astronautical Congress (IAC) in 2011, Southafrica.

Appendix A: Transformation Routines

This material is adapted from [12][13].

- Classical Orbital Elements to Position and Velocity in ECI Coordinate frame.
 - Given orbital elements: $a, e, i, \Omega, \omega, \nu$
 - Convert into Perifocal Coordinate frame

$$p = a(1 - e^2)$$

$$\vec{r}_{perifocal} = \begin{bmatrix} \frac{p \cos \nu}{1 + e \cos \nu} \\ \frac{p \sin \nu}{1 + e \cos \nu} \\ 0 \end{bmatrix} \quad \vec{v}_{perifocal} = \begin{bmatrix} -\sqrt{\frac{\mu_E}{p}} \sin \nu \\ \sqrt{\frac{\mu_E}{p}} (e + \cos \nu) \\ 0 \end{bmatrix}$$

- Compute rotation matrices

$$\mathbf{PQW} = \mathfrak{R}_z(-\Omega) * \mathfrak{R}_x(-i) * \mathfrak{R}_z(-\omega)$$

where, $\mathfrak{R}_z(.)$ and $\mathfrak{R}_x(.)$ are rotation matrices for rotation about “z” and “x” axes respectively. For example for θ = rotation angle the individual rotation matrices can be computed as:

$$\mathfrak{R}_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & +\cos \theta & +\sin \theta \\ 0 & -\sin \theta & +\cos \theta \end{bmatrix}, \quad \mathfrak{R}_z(\theta) = \begin{bmatrix} +\cos \theta & +\sin \theta & 0 \\ -\sin \theta & +\cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Three dimensional ECI position and velocity is obtained as:

$$\vec{r} = \mathbf{PQW} * \vec{r}_{perifocal}$$

$$\vec{v} = \mathbf{PQW} * \vec{v}_{perifocal}$$

- From Position and Velocity in ECI coordinate frame to classical orbital elements.
 - Given Position and Velocity

$$\vec{r}, \vec{v}$$

- Compute orbital elements

$$\vec{h} = \vec{r} \times \vec{v} \quad h = |\vec{h}|$$

$$\vec{n} = \hat{K} \times \vec{h}$$

where, $\hat{K} = (0,0,1)^T$

$$\vec{e} = \frac{\left(v^2 - \frac{\mu_E}{r}\right) \vec{r} - (\vec{r} \cdot \vec{v}) \vec{v}}{\mu_E}, \quad e = |\vec{e}|$$

$$\varepsilon = \frac{v^2}{2} - \frac{\mu_E}{r}$$

$$a = -\frac{\mu_E}{2\varepsilon}$$

where, $r = |\vec{r}|$, $v = |\vec{v}|$

$$i = \cos^{-1} \left(\frac{h_z}{h} \right)$$

$$\Omega = \cos^{-1} \left(\frac{n_x}{|\vec{n}|} \right), \text{ if } (n_y < 0) \text{ then } \Omega = 360^\circ - \Omega$$

$$\omega = \cos^{-1} \left(\frac{\vec{n} \cdot \vec{e}}{|\vec{n}| |\vec{e}|} \right), \text{ if } (e_z < 0) \text{ then } \omega = 360^\circ - \omega$$

$$\nu = \cos^{-1} \left(\frac{\vec{e} \cdot \vec{r}}{|\vec{e}| |\vec{r}|} \right), \text{ if } (\vec{r} \cdot \vec{v} < 0) \text{ then } \nu = 360^\circ - \nu$$

where, h_z = Z-axis component of \vec{h} , n_x, n_y = X and Y axis component of \vec{n} and e_z = Z-axis component of \vec{e}

Appendix B: Partial's for State Transition Matrix Kepler's Equation

This material is adapted from [13].

$$\frac{\partial \mathbf{y}(t)}{\partial \mathbf{y}(t_0)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{\partial M(t)}{\partial a(t_0)} & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\frac{\partial M(t)}{\partial a(t_0)} = -\frac{3n}{2a}(t - t_0)$$

$$\frac{\partial \mathbf{P}(\mathbf{x}_0, t)}{\partial \mathbf{y}(t)} = \begin{bmatrix} \frac{\partial \vec{r}}{\partial \mathbf{y}} & \frac{\partial \vec{v}}{\partial \mathbf{y}} \end{bmatrix}^T$$

where,

$$\frac{\partial \vec{r}}{\partial \mathbf{y}} = \frac{\partial \hat{X}}{\partial \mathbf{y}} \mathbf{P} + \frac{\partial \hat{Y}}{\partial \mathbf{y}} \mathbf{Q}, \quad \frac{\partial \vec{v}}{\partial \mathbf{y}} = \frac{\partial \dot{\hat{X}}}{\partial \mathbf{y}} \mathbf{P} + \frac{\partial \dot{\hat{Y}}}{\partial \mathbf{y}} \mathbf{Q}$$

for $\mathbf{y} = a, e, M$

$$\frac{\partial \vec{r}}{\partial \mathbf{y}} = \hat{X} \frac{\partial \mathbf{P}}{\partial \mathbf{y}} + \hat{Y} \frac{\partial \mathbf{Q}}{\partial \mathbf{y}}, \quad \frac{\partial \vec{v}}{\partial \mathbf{y}} = \dot{\hat{X}} \frac{\partial \mathbf{P}}{\partial \mathbf{y}} + \dot{\hat{Y}} \frac{\partial \mathbf{Q}}{\partial \mathbf{y}}$$

for $\mathbf{y} = \Omega, \omega, i$

$$\frac{\partial (\hat{X}, \hat{Y})^T}{\partial (a, e, M)^T} = \begin{bmatrix} \frac{\hat{X}}{a} & \left(-a - \frac{\hat{Y}^2}{r(1-e^2)} \right) & \frac{\dot{\hat{X}}}{n} \\ \frac{\hat{Y}}{a} & \left(\frac{\hat{X}\hat{Y}}{r(1-e^2)} \right) & \frac{\dot{\hat{Y}}}{n} \end{bmatrix}$$

$$\frac{\partial (\dot{\hat{X}}, \dot{\hat{Y}})^T}{\partial (a, e, M)^T} = \begin{bmatrix} \frac{\dot{\hat{X}}}{2a} & \dot{\hat{X}} \left(\frac{a}{r} \right)^2 \left(2 \left(\frac{\hat{X}}{a} \right) + \frac{e}{1-e^2} \left(\frac{\hat{Y}}{a} \right)^2 \right) & -n \left(\frac{a}{r} \right)^3 \hat{X} \\ \frac{\dot{\hat{Y}}}{2a} & \frac{n}{\sqrt{1-e^2}} \left(\frac{a}{r} \right)^2 \left(\frac{\hat{X}^2}{r} - \frac{\hat{Y}^2}{a(1-e^2)} \right) & -n \left(\frac{a}{r} \right)^3 \hat{Y} \end{bmatrix}$$

$$\frac{\partial \mathbf{P}}{\partial i} = \sin \omega \cdot \mathbf{W}, \quad \frac{\partial \mathbf{Q}}{\partial i} = \cos \omega \cdot \mathbf{W}$$

$$\frac{\partial \mathbf{P}}{\partial \Omega} = \begin{bmatrix} -P_y \\ +P_x \\ 0 \end{bmatrix}, \quad \frac{\partial \mathbf{Q}}{\partial \Omega} = \begin{bmatrix} -Q_y \\ +Q_x \\ 0 \end{bmatrix}$$

$$\frac{\partial \mathbf{P}}{\partial \omega} = +\mathbf{Q}, \quad \frac{\partial \mathbf{Q}}{\partial \omega} = -\mathbf{P}$$

where, P_x, P_y, Q_x and Q_y are x and y component of \mathbf{P} and \mathbf{Q}

$$\left(\frac{\partial \mathbf{x}_0}{\partial \mathbf{y}(t_0)} \right)^{-1} = \left(\frac{\partial \mathbf{y}}{\partial \mathbf{x}} \right) \Big|_{\substack{\mathbf{y}=\mathbf{y}(t_0) \\ \mathbf{x}=\mathbf{x}_0}}$$

where,

$$\begin{pmatrix} \frac{\partial \mathbf{y}}{\partial \vec{r}} & \frac{\partial \mathbf{y}}{\partial \vec{v}} \end{pmatrix} = P(\mathbf{y}, \mathbf{y}) \begin{pmatrix} + \left(\frac{\partial \vec{v}}{\partial \mathbf{y}} \right)^T & - \left(\frac{\partial \vec{r}}{\partial \mathbf{y}} \right)^T \end{pmatrix}$$

where, P = anti-symmetric 6×6 matrix made up of Poisson parentheses:

$$P(y_i, y_j) = \left(\frac{\partial y_i}{\partial \vec{r}} \right) \cdot \left(\frac{\partial y_j}{\partial \vec{v}} \right)^T - \left(\frac{\partial y_j}{\partial \vec{r}} \right) \cdot \left(\frac{\partial y_i}{\partial \vec{v}} \right)^T$$

with following independent matrix elements:

$$\begin{aligned} P(a, M) &= -P(M, a) = \frac{-2}{na} \\ P(e, \omega) &= -P(\omega, e) = \frac{\sqrt{1-e^2}}{na^2 e} \\ P(e, M) &= -P(M, e) = \frac{-(1-e^2)}{na^2 e} \\ P(i, \Omega) &= -P(\Omega, i) = \frac{1}{na^2 \sqrt{1-e^2} \sin i} \\ P(i, \omega) &= -P(\omega, i) = \frac{-\cot i}{na^2 \sqrt{1-e^2}} \end{aligned}$$

rest of the Poisson parentheses vanish.

Appendix C: Epicycle Coefficients for Geopotential Zonal Harmonic Terms up to J_4

This material has been adapted from [2]:

J_2 Secular Terms are:

$$\begin{aligned}\rho_2 &= -\frac{1}{4} A_2 (3 \cos^2 I_0 - 1) \\ \mathcal{G}_2 &= -\frac{3}{2} A_2 \cos I_0 \\ \kappa_2 &= \frac{3}{4} A_2 (5 \cos^2 I_0 - 1)\end{aligned}$$

J_2 coefficients for short periodic terms are:

$$\begin{aligned}\Delta r_2 &= \frac{1}{4} A_2 \sin^2 I_0 \\ \Delta I_2 &= -\frac{3}{8} A_2 \sin 2I_0 \\ \Delta \Omega_2 &= \frac{3}{4} A_2 \cos I_0 \\ \Delta \lambda_2 &= -\frac{1}{8} A_2 (6 - 7 \sin^2 I_0)\end{aligned}$$

The J_3 coefficients for the long-periodic terms using only J_2 in κ_2 are:

$$\chi_3 = \frac{J_3}{2J_2} \left(\frac{R_E}{a} \right) \sin I_0$$

and short periodic terms are:

$$\begin{aligned}\Delta I_3^0 &= -\frac{3}{8} J_3 \left(\frac{R_E}{a} \right)^3 \cos I_0 (4 - 5 \sin^2 I_0) \\ \Delta \Omega_3^0 &= \frac{3}{8} J_3 \left(\frac{R_E}{a} \right)^3 \cot I_0 (4 - 15 \sin^2 I_0) \\ \Delta \lambda_3^0 &= \frac{3}{8} J_3 \left(\frac{R_E}{a} \right)^3 \sin I_0 [(4 - 5 \sin^2 I_0) - (4 - 15 \sin^2 I_0) \cot^2 I_0]\end{aligned}$$

and

$$\Delta r_3^1 = \frac{5}{32} J_3 \left(\frac{R_E}{a} \right)^3 \sin^3 I_0$$

$$\Delta I_3^1 = \frac{5}{8} J_3 \left(\frac{R_E}{a} \right)^3 \cos I_0 (\sin^2 I_0)$$

$$\Delta \Omega_3^1 = \frac{5}{8} J_3 \left(\frac{R_E}{a} \right)^3 \cot I_0 (\sin^2 I_0)$$

$$\Delta \lambda_3^1 = -\frac{5}{48} J_3 \left(\frac{R_E}{a} \right)^3 \sin I_0 [(6 - 7 \sin^2 I_0)]$$

The J_4 coefficients for secular terms are:

$$\rho_4 = \frac{9}{64} J_4 \left(\frac{R_E}{a} \right)^4 (8 - 40 \sin^2 I_0 + 35 \sin^4 I_0)$$

$$\vartheta_4 = \frac{15}{16} J_4 \left(\frac{R_E}{a} \right)^4 \cos I_0 (4 - 7 \sin^2 I_0)$$

$$\kappa_4 = -\frac{3}{64} J_4 \left(\frac{R_E}{a} \right)^4 (136 - 500 \sin^2 I_0 + 385 \sin^4 I_0)$$

and for short-periodic terms are:

$$\Delta r_4^1 = -\frac{5}{16} J_4 \left(\frac{R_E}{a} \right)^4 \sin^2 I_0 (6 - 7 \sin^2 I_0)$$

$$\Delta I_4^1 = \frac{5}{32} J_4 \left(\frac{R_E}{a} \right)^4 \sin 2I_0 (6 - 7 \sin^2 I_0)$$

$$\Delta \Omega_4^1 = -\frac{5}{8} J_4 \left(\frac{R_E}{a} \right)^4 \cos I_0 (3 - 7 \sin^2 I_0)$$

$$\Delta \lambda_4^1 = \frac{5}{32} J_4 \left(\frac{R_E}{a} \right)^4 (12 - 34 \sin^2 I_0 + 21 \sin^4 I_0)$$

$$\Delta r_4^2 = -\frac{7}{64} J_4 \left(\frac{R_E}{a} \right)^4 \sin^4 I_0$$

$$\Delta I_4^2 = \frac{35}{128} J_4 \left(\frac{R_E}{a} \right)^4 \sin 2I_0 \sin^2 I_0$$

$$\Delta \Omega_4^2 = -\frac{35}{64} J_4 \left(\frac{R_E}{a} \right)^4 \cos I_0 \sin^2 I_0$$

$$\Delta \lambda_4^2 = \frac{7}{256} J_4 \left(\frac{R_E}{a} \right)^4 \sin^2 I_0 (20 - 23 \sin^2 I_0)$$

J_2^2 coefficients for secular terms are:

$$\rho_{2s} = -\frac{1}{32} A_2^2 (16 + 24 \sin^2 I_0 - 49 \sin^4 I_0)$$

$$\vartheta_{2s} = \frac{3}{8} A_2^2 \cos I_0 (11 - 20 \sin^2 I_0)$$

$$\kappa_{2s} = \frac{3}{16} A_2^2 (14 + 17 \sin^2 I_0 - 35 \sin^4 I_0)$$

J_2^2 2α short periodic terms are:

$$\Delta r_{2s}^1 = -\frac{1}{4} A_2^2 \sin^2 I_0 (9 - 10 \sin^2 I_0)$$

$$\Delta I_{2s}^1 = \frac{3}{32} A_2^2 \sin 2I_0 (11 - 15 \sin^2 I_0)$$

$$\Delta \Omega_{2s}^1 = -\frac{3}{16} A_2^2 \cos I_0 (8 - 21 \sin^2 I_0)$$

$$\Delta \lambda_{2s}^1 = \frac{1}{32} A_2^2 (48 - 190 \sin^2 I_0 + 139 \sin^4 I_0)$$

J_2^2 4α short periodic terms are:

$$\Delta r_{2s}^2 = -\frac{1}{32} A_2^2 \sin^4 I_0$$

$$\Delta I_{2s}^2 = \frac{3}{128} A_2^2 \sin 2I_0 (3 + 5 \sin^2 I_0)$$

$$\Delta \Omega_{2s}^2 = -\frac{3}{32} A_2^2 \cos I_0 (3 + \sin^2 I_0)$$

$$\Delta \lambda_{2s}^2 = \frac{1}{64} A_2^2 (18 - 3 \sin^2 I_0 - 17 \sin^4 I_0)$$

where,

$$A_2 = J_2 \left(\frac{R_E}{a} \right)^2$$

$$J_2 = 1.08262668355 \times 10^{-3}$$

$$J_3 = -2.53265648533 \times 10^{-6}$$

$$J_4 = -1.61962159137 \times 10^{-6}$$

Equatorial radius of Earth

$$R_E = 6378.136 \text{ km}$$

Appendix D: Partial for Epicyclic Orbit Analysis

This material has been adapted from [14]:

Matrix of partials $\frac{\partial \mathbf{P}(\mathbf{x}_0, t)}{\partial \mathbf{y}(t)}$ for Epicycle Orbit Estimation Problem:

$$\frac{\partial(X, Y, Z)}{\partial(r, I, \Omega, \lambda, \dot{r}, \dot{I}, \dot{\Omega}, \dot{\lambda})} = \begin{bmatrix} f_1 & rg_4 \sin \Omega & -rf_3 & -rf_2 & 0 & 0 & 0 & 0 \\ f_3 & -rg_4 \cos \Omega & rf_1 & -rf_4 & 0 & 0 & 0 & 0 \\ g_4 & rg_2 & 0 & rg_3 & 0 & 0 & 0 & 0 \end{bmatrix}$$

And

$$\frac{\partial \dot{X}}{\partial(r, I, \Omega, \lambda, \dot{r}, \dot{I}, \dot{\Omega}, \dot{\lambda})} = \begin{bmatrix} \dot{I}g_4 \sin \Omega - \dot{\Omega}f_3 - \dot{\lambda}f_2 \\ \dot{r}g_4 \sin \Omega + r(\dot{I}g_2 \sin \Omega + \dot{\Omega}g_4 \cos \Omega + \dot{\lambda}g_3 \sin \Omega) \\ -\dot{r}f_3 + r(\dot{I}g_4 \cos \Omega - \dot{\Omega}f_1 + \dot{\lambda}f_4) \\ -\dot{r}f_2 + r(\dot{I}g_3 \sin \Omega + \dot{\Omega}f_4 - \dot{\lambda}f_1) \\ f_1 \\ rg_4 \sin \Omega \\ -rf_3 \\ -rf_2 \end{bmatrix}^T$$

$$\frac{\partial \dot{Y}}{\partial(r, I, \Omega, \lambda, \dot{r}, \dot{I}, \dot{\Omega}, \dot{\lambda})} = \begin{bmatrix} -\dot{I}g_4 \cos \Omega + \dot{\Omega}f_1 - \dot{\lambda}f_4 \\ -\dot{r}g_4 \cos \Omega + r(\dot{I}g_2 \cos \Omega - \dot{\Omega}g_4 \sin \Omega + \dot{\lambda}g_3 \cos \Omega) \\ \dot{r}f_1 + r(\dot{I}g_4 \sin \Omega - \dot{\Omega}f_3 - \dot{\lambda}f_2) \\ -\dot{r}f_4 - r(\dot{I}g_3 \cos \Omega + \dot{\Omega}f_2 + \dot{\lambda}f_3) \\ f_3 \\ -rg_4 \cos \Omega \\ rf_1 \\ -rf_4 \end{bmatrix}^T$$

$$\frac{\partial \dot{Z}}{\partial(r, I, \Omega, \lambda, \dot{r}, \dot{I}, \dot{\Omega}, \dot{\lambda})} = \begin{bmatrix} \dot{I}g_2 + \dot{\lambda}g_3 \\ \dot{r}g_2 - r(\dot{I}g_4 - \dot{\lambda}g_1) \\ 0 \\ \dot{r}g_3 + r(\dot{I}g_1 - \dot{\lambda}g_4) \\ g_4 \\ rg_2 \\ 0 \\ rg_3 \end{bmatrix}^T$$

The second partial derivative matrix $\frac{\partial \mathbf{y}(t)}{\partial \mathbf{x}_0}$ is found through Equation: 3.43 and 3.45. We will use the

abbreviation of $\partial_x = \frac{\partial}{\partial x}$, the partial derivative with respect to positional coordinates are obtained as follows:

$$\frac{1}{a} \frac{\partial r}{\partial(a, \xi_p, \eta_p, I_0, \Omega_0, \alpha_0)} = \begin{bmatrix} \frac{r}{a^2} + \partial_a \rho + (\partial_a \Delta r_2) \cos 2\beta + (\xi_p \sin \alpha - \eta_p \cos \alpha)(\partial_a \alpha) \\ -(2\Delta r_2 \sin 2\beta)(\partial_a \beta) \\ -\cos \alpha \\ -\sin \alpha \\ \partial_{I_0} \rho + (\partial_{I_0} \Delta r_2) \cos 2\beta - (2\Delta r_2 \sin 2\beta) \partial_{I_0} \beta \\ 0 \\ (\xi_p \sin \alpha - \eta_p \cos \alpha) - (1 + \kappa) 2\Delta r_2 \sin 2\beta \end{bmatrix}^T$$

where,

$$\partial_a \alpha = \frac{\alpha}{n} (\partial_a n) = -\frac{3}{2} \frac{\alpha}{a}$$

$$\partial_a \beta = (1 + \kappa) \partial_a \alpha + (\partial_a \kappa) \alpha = \left(-\frac{3(1 + \kappa)}{2a} + \partial_a \kappa \right) \alpha$$

$$\partial_{I_0} = (\partial_{I_0} \kappa) \alpha$$

$$\frac{\partial I}{\partial(a, \xi_p, \eta_p, I_0, \Omega_0, \alpha_0)} = \begin{bmatrix} \partial_a \Delta I_2 (1 - \cos 2\beta) + 2\Delta I_2 \sin 2\beta (\partial_a \beta) \\ 0 \\ 0 \\ 1 + (\partial_{I_0} \Delta I_2) (1 - \cos 2\beta) + 2\Delta I_2 \sin 2\beta (\partial_{I_0} \beta) \\ 0 \\ 2(1 + \kappa) \Delta I_2 \sin 2\beta \end{bmatrix}^T$$

$$\frac{\partial \Omega}{\partial(a, \xi_p, \eta_p, I_0, \Omega_0, \alpha_0)} = \begin{bmatrix} \partial_a \mathcal{G} + \mathcal{G}(\partial_a \alpha) + (\partial_a \Delta \Omega_2) \sin 2\beta + 2\Delta \Omega_2 \cos 2\beta (\partial_a \beta) \\ 0 \\ 0 \\ (\partial_{I_0} \mathcal{G})\alpha + (\partial_{I_0} \Delta \Omega_2)(\sin 2\beta) + 2\Delta \Omega_2 \cos 2\beta (\partial_{I_0} \beta) \\ 1 \\ \mathcal{G} + 2(1 + \kappa)\Delta \Omega_2 \cos 2\beta \end{bmatrix}^T$$

$$\frac{\partial \lambda}{\partial(a, \xi_p, \eta_p, I_0, \Omega_0, \alpha_0)} = \begin{bmatrix} \partial_a \beta + (\partial_a \Delta \lambda_2) \sin 2\beta + 2(\xi_p \cos \alpha + \eta_p \sin \alpha)(\partial_a \alpha) \\ + 2\Delta \lambda_2 \cos 2\beta (\partial_a \beta) \\ 2 \sin \alpha \\ 2(1 - \cos \alpha) \\ (\partial_{I_0} \kappa)\alpha + (\partial_{I_0} \Delta \lambda_2)(\sin 2\beta) + 2\Delta \lambda_2 \cos 2\beta (\partial_{I_0} \beta) \\ 0 \\ (1 + \kappa) + 2(\xi_p \cos \alpha + \eta_p \sin \alpha) + 2(1 + \kappa)\Delta \lambda_2 \cos 2\beta \end{bmatrix}^T$$

And the partial derivatives with respect to velocity coordinates are:

$$\frac{1}{an} \frac{\partial \dot{r}}{\partial(a, \xi_p, \eta_p, I_0, \Omega_0, \alpha_0)} = \begin{bmatrix} \frac{\dot{r}}{2a^2 n} - (1 + \kappa)2(\partial_a \Delta r_2) \sin 2\beta - 2\Delta r_2 \sin 2\beta (\partial_a \kappa) \\ + (\xi_p \cos \alpha + \eta_p \sin \alpha)(\partial_a \alpha) \\ - (1 + \kappa)(4\Delta r_2 \sin 2\beta)(\partial_a \beta) \\ \sin \alpha \\ - \cos \alpha \\ - (1 + \kappa)2(\partial_{I_0} \Delta r_2) \sin 2\beta - 2\Delta r_2 \sin 2\beta (\partial_{I_0} \kappa) \\ - (1 + \kappa)4\Delta r_2 \sin 2\beta (\partial_{I_0} \beta) \\ 0 \\ (\xi_p \cos \alpha + \eta_p \sin \alpha) - (1 + \kappa)^2 (4\Delta r_2 \cos 2\beta) \end{bmatrix}^T$$

$$\begin{aligned}
\frac{1}{n} \frac{\partial \dot{I}}{\partial(a, \xi_p, \eta_p, I_0, \Omega_0, \alpha_0)} &= \begin{bmatrix} 2(1+\kappa)[(\partial_a \Delta I_2) \sin 2\beta + 2\Delta I_2 \cos 2\beta (\partial_a \beta)] \\ + 2\Delta I_2 \sin 2\beta (\partial_a \kappa) - \left(\frac{3}{2}\right) \frac{\dot{I}}{an} \\ 0 \\ 0 \\ 2(1+\kappa)[(\partial_{I_0} \Delta I_2) \sin 2\beta + 2\Delta I_2 \cos 2\beta (\partial_{I_0} \beta)] \\ + 2\Delta I_2 \sin 2\beta (\partial_{I_0} \kappa) \\ 0 \\ 4(1+\kappa)^2 \Delta I_2 \cos 2\beta \end{bmatrix}^T \\
\\
\frac{1}{n} \frac{\partial \dot{\Omega}}{\partial(a, \xi_p, \eta_p, I_0, \Omega_0, \alpha_0)} &= \begin{bmatrix} \partial_a \mathcal{G} + 2(1+\kappa)(\partial_a \Delta \Omega_2) \cos 2\beta + 2\Delta \Omega_2 \cos 2\beta (\partial_a \kappa) \\ - 2(1+\kappa) \Delta \Omega_2 \sin 2\beta (\partial_a \beta) - \left(\frac{3}{2}\right) \frac{\dot{\Omega}}{an} \\ 0 \\ 0 \\ \partial_{I_0} \mathcal{G} + 2(1+\kappa)[(\partial_{I_0} \Delta \Omega_2) \cos 2\beta - 2\Delta \Omega_2 \sin 2\beta (\partial_{I_0} \beta)] \\ + 2\Delta \Omega_2 \cos 2\beta (\partial_{I_0} \kappa) \\ 0 \\ - 4(1+\kappa)^2 \Delta \Omega_2 \cos 2\beta \end{bmatrix}^T \\
\\
\frac{1}{n} \frac{\partial \dot{\lambda}}{\partial(a, \xi_p, \eta_p, I_0, \Omega_0, \alpha_0)} &= \begin{bmatrix} \partial_a \kappa + 2(1+\kappa)[(\partial_a \Delta \lambda_2) \cos 2\beta] + 2\Delta \lambda_2 \cos 2\beta (\partial_a \kappa) \\ - 2(\xi_p \sin \alpha - \eta_p \cos \alpha)(\partial_a \alpha) - 2(1+\kappa)(2\Delta \lambda_2 \sin 2\beta)(\partial_a \beta) \\ - \left(\frac{3}{2}\right) \frac{\dot{\lambda}}{an} \\ 2 \cos \alpha \\ 2 \sin \alpha \\ \partial_{I_0} \kappa + 2(1+\kappa)[(\partial_{I_0} \Delta \lambda_2) \cos 2\beta] + 2\Delta \lambda_2 \cos 2\beta (\partial_{I_0} \kappa) \\ - 2(1+\kappa)(2\Delta \lambda_2 \sin 2\beta)(\partial_{I_0} \beta) \\ 0 \\ - 2(\xi_p \sin \alpha - \eta_p \cos \alpha) - 2(1+\kappa)^2 (2\Delta \lambda_2 \sin 2\beta) \end{bmatrix}^T
\end{aligned}$$

Partial Derivatives of the J_2 first order secular perturbations are:

$$\begin{aligned}
\frac{\partial(\rho_2, \mathcal{G}_2, \kappa_2)}{\partial a} &= \begin{bmatrix} \frac{1}{2} \frac{A_2}{a} (3 \cos^2 I_0 - 1) \\ \frac{3A_2}{a} \cos I_0 \\ -\frac{3}{2} \frac{A_2}{a} (5 \cos^2 I_0 - 1) \end{bmatrix} & \frac{\partial(\rho_2, \mathcal{G}_2, \kappa_2)}{\partial I_0} &= \begin{bmatrix} \frac{3}{4} A_2 \sin 2I_0 \\ \frac{3}{2} A_2 \sin I_0 \\ -\frac{15}{4} A_2 \sin 2I_0 \end{bmatrix}
\end{aligned}$$

Partial derivatives of J_2 first order short periodic perturbations are:

$$\frac{\partial(\Delta r_2, \Delta I_2, \Delta \Omega_2, \Delta \lambda_2)}{\partial a} = \begin{bmatrix} -\frac{1}{2} \frac{A_2}{a} \sin^2 I_0 \\ \frac{3}{4} \frac{A_2}{a} \sin 2I_0 \\ -\frac{3}{2} \frac{A_2}{a} \cos I_0 \\ \frac{1}{4} \frac{A_2}{a} (7 \cos^2 I_0 - 1) \end{bmatrix}$$

$$\frac{\partial(\Delta r_2, \Delta I_2, \Delta \Omega_2, \Delta \lambda_2)}{\partial I_0} = \begin{bmatrix} \frac{1}{4} A_2 \sin 2I_0 \\ -\frac{3}{4} A_2 \cos 2I_0 \\ -\frac{3}{4} A_2 \sin I_0 \\ \frac{7}{8} A_2 \sin 2I_0 \end{bmatrix}$$

where,

$$A_2 = J_2 \left(\frac{R}{a} \right)^2$$

Appendix E: Analytical Solution of Modified HCW Equations by Schweighart and Sedwick

The solutions of modified HCW by SS are expressed as [20]:

$$\begin{aligned}
 x(t) &= \left(\frac{4(1+s) - (3+5s)\cos(nt\sqrt{1-s})}{1-s} \right) x_0 + \frac{\sin(nt\sqrt{1-s})}{n\sqrt{1-s}} \dot{x}_0 + \frac{2\sqrt{1+s}}{n(1-s)} (1 - \cos(nt\sqrt{1-s})) \dot{y}_0 \\
 y(t) &= y_0 + \frac{2\sqrt{1+s}}{n(1-s)} (\cos(nt\sqrt{1-s}) - 1) \dot{x}_0 + \frac{2\sqrt{1+s}(3+5s)}{1-s} \left(\frac{\sin(nt\sqrt{1-s})}{\sqrt{1-s}} - nt \right) x_0 \\
 &\quad + \left(t - \frac{4(1+s)}{1-s} t + \frac{4(1+s)}{n(\sqrt{1-s})^3} \sin(nt\sqrt{1-s}) \right) \dot{y}_0 \\
 z(t) &= (lt + m) \sin(qt + \phi) \\
 \dot{x}(t) &= \cos(nt\sqrt{1-s}) \dot{x}_0 + \frac{n(3+5s)}{\sqrt{1-s}} \sin(nt\sqrt{1-s}) x_0 + \frac{2\sqrt{1+s}}{(1-s)} \sin(nt\sqrt{1-s}) \dot{y}_0 \\
 \dot{y}(t) &= \frac{2n\sqrt{1+s}(3+5s)}{1-s} (\cos(nt\sqrt{1-s}) - 1) x_0 - \frac{2\sqrt{1+s}}{\sqrt{1-s}} \sin(nt\sqrt{1-s}) \dot{x}_0 \\
 &\quad + \left(\frac{4(1+s)}{1-s} \cos(nt\sqrt{1-s}) + 1 - \frac{4(1+s)}{1-s} \right) \dot{y}_0 \\
 \dot{z}(t) &= l \sin(qt + \phi) + q(lt + m) \cos(qt + \phi)
 \end{aligned}$$

Two initial velocity conditions are specified to remove the offset and drift in in-plane motion:

$$\dot{x}_0 = \frac{ny_0}{2} \frac{(1-s)}{\sqrt{1+s}}, \dot{y}_0 = -2nx_0\sqrt{1+s}$$

For clarity all of the constants needed for these equations are given as:

$$\begin{aligned}
 s &= \frac{3J_2 R_E^2}{8r_c^2} (1 + 3\cos 2i_c), c = \sqrt{1+s}, n = \sqrt{\frac{\mu_E}{r_c^3}}, k = nc + \frac{3nJ_2 R_E^2}{2r_c^2} \cos^2 i_c, \Delta\Omega_0 = \frac{z_0}{r_c \sin i_c} \\
 \gamma_0 &= \cot^{-1} \left[\frac{\cot i_c \sin i_d - \cos i_d \cos \Delta\Omega_0}{\sin \Delta\Omega_0} \right], \Phi_0 = \cos^{-1} [\cos i_d \cos i_c + \sin i_d \sin i_c \cos \Delta\Omega_0]
 \end{aligned}$$

$$\dot{\Omega}_d = -\frac{3nJ_2R_E^2}{2r_c^2}\cos i_d, \dot{\Omega}_c = -\frac{3nJ_2R_E^2}{2r_c^2}\cos i_c, q = nc - (\cos \gamma_0 \sin \gamma_0 \cot \Delta\Omega_0 - \sin^2 \gamma_0 \cos i_d)$$

$$(\dot{\Omega}_d - \dot{\Omega}_c) - \dot{\Omega}_d \cos i_d, l = -r_c \frac{\sin i_d \sin i_c \sin \Delta\Omega_0}{\sin \Phi_0} (\dot{\Omega}_d - \dot{\Omega}_c)$$

$$m \sin \phi = z_0, l \sin \phi + qm \cos \phi = \dot{z}_0$$

$$i_d = \frac{\dot{z}_0}{kr_c} + i_c$$

R_E , Equatorial Radius of Earth, μ_E , Earth Gravitational Parameter and J_2 , Geo-potential Second Zonal Harmonic. Subscripts c and d are for chief and deputy satellites respectively.

Appendix F: Partial for Schweighart and Sedwick J_2 Modified HCW Equations

The partials of ECI position coordinates with respect to epoch relative state vector is expressed as:

$$\frac{\partial \mathbf{X}_d}{\partial \mathbf{x}_0} = \begin{bmatrix} \hat{e}_{RX} \left\{ 1 + \frac{(3+5s)}{1-s} - \frac{(3+5s)}{1-s} \cos(nt\sqrt{1-s}) \right\} + \hat{e}_{IX} \left\{ \frac{2\sqrt{1+s}(3+5s)}{\sqrt[3]{1-s}} \sin \cos(nt\sqrt{1-s}) \right. \\ \left. - \frac{2n\sqrt{1+s}(3+5s)}{1-s} t \right\} \\ \hat{e}_{CX} \left\{ \frac{\partial z}{\partial z_0} \right\} \\ \hat{e}_{RX} \left\{ \frac{\sin(nt\sqrt{1-s})}{n\sqrt{1-s}} \right\} + \hat{e}_{IX} \left\{ \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) - \frac{2\sqrt{1+s}}{n(1-s)} \right\} \\ \hat{e}_{RX} \left\{ \frac{2\sqrt{1+s}}{n(1-s)} - \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) \right\} + \hat{e}_{IX} \left\{ \frac{4(1+s)}{n(\sqrt[3]{1-s})} \sin \cos(nt\sqrt{1-s}) + t \right. \\ \left. - \frac{4(1+s)}{1-s} t \right\} \\ \hat{e}_{CX} \left\{ \frac{\partial z}{\partial z_0} \right\} \end{bmatrix}^T$$

$$\begin{aligned}
\frac{\partial Y_d}{\partial \mathbf{x}_0} = & \left[\begin{array}{c} \hat{e}_{RY} \left\{ 1 + \frac{(3+5s)}{1-s} - \frac{(3+5s)}{1-s} \cos(nt\sqrt{1-s}) \right\} + \hat{e}_{IY} \left\{ \frac{2\sqrt{1+s}(3+5s)}{\sqrt[3]{1-s}} \sin \cos(nt\sqrt{1-s}) \right. \\ \left. - \frac{2n\sqrt{1+s}(3+5s)}{1-s} t \right\} \\ \hat{e}_{CY} \left\{ \frac{\partial z}{\partial z_0} \right\} \\ \hat{e}_{RY} \left\{ \frac{\sin(nt\sqrt{1-s})}{n\sqrt{1-s}} \right\} + \hat{e}_{IY} \left\{ \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) - \frac{2\sqrt{1+s}}{n(1-s)} \right\} \\ \hat{e}_{RY} \left\{ \frac{2\sqrt{1+s}}{n(1-s)} - \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) \right\} + \hat{e}_{IY} \left\{ \frac{4(1+s)}{n(\sqrt[3]{1-s})} \sin \cos(nt\sqrt{1-s}) + t \right. \\ \left. - \frac{4(1+s)}{1-s} t \right\} \\ \hat{e}_{CY} \left\{ \frac{\partial z}{\partial z_0} \right\} \end{array} \right]^T \\
\\
\frac{\partial Z_d}{\partial \mathbf{x}_0} = & \left[\begin{array}{c} \hat{e}_{RZ} \left\{ 1 + \frac{(3+5s)}{1-s} - \frac{(3+5s)}{1-s} \cos(nt\sqrt{1-s}) \right\} + \hat{e}_{IZ} \left\{ \frac{2\sqrt{1+s}(3+5s)}{\sqrt[3]{1-s}} \sin \cos(nt\sqrt{1-s}) \right. \\ \left. - \frac{2n\sqrt{1+s}(3+5s)}{1-s} t \right\} \\ \hat{e}_{CZ} \left\{ \frac{\partial z}{\partial z_0} \right\} \\ \hat{e}_{RZ} \left\{ \frac{\sin(nt\sqrt{1-s})}{n\sqrt{1-s}} \right\} + \hat{e}_{IZ} \left\{ \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) - \frac{2\sqrt{1+s}}{n(1-s)} \right\} \\ \hat{e}_{RZ} \left\{ \frac{2\sqrt{1+s}}{n(1-s)} - \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) \right\} + \hat{e}_{IZ} \left\{ \frac{4(1+s)}{n(\sqrt[3]{1-s})} \sin \cos(nt\sqrt{1-s}) + t \right. \\ \left. - \frac{4(1+s)}{1-s} t \right\} \\ \hat{e}_{CZ} \left\{ \frac{\partial z}{\partial z_0} \right\} \end{array} \right]^T
\end{aligned}$$

The partials for the ECI velocity coordinates with respect to initial relative state vector are:-

$$\frac{\partial \dot{X}_d}{\partial \mathbf{x}_0} = \begin{bmatrix} \frac{n(3+5s)}{\sqrt{1-s}} \sin(nt\sqrt{1-s}) - \mathbf{A}_3 \left\{ \frac{2\sqrt{1+s}(3+5s)}{\sqrt[3]{1-s}} \sin(nt\sqrt{1-s}) - \frac{2n\sqrt{1+s}(3+5s)}{1-s} t \right\} \\ -\mathbf{A}_3 \\ \mathbf{A}_2 \left\{ \frac{\partial z}{\partial z_0} \right\} \\ \cos(nt\sqrt{1-s}) - \mathbf{A}_3 \left\{ \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) - \frac{2\sqrt{1+s}}{n(1-s)} \right\} \\ \frac{2\sqrt{1+s}}{\sqrt{1-s}} \sin(nt\sqrt{1-s}) - \mathbf{A}_3 \left\{ \frac{4(1+s)}{n(\sqrt[3]{1-s})} \sin(nt\sqrt{1-s}) + t - \frac{4(1+s)}{1-s} t \right\} \\ \mathbf{A}_2 \left\{ \frac{\partial z}{\partial z_0} \right\} \end{bmatrix}^T$$

$$\frac{\partial \dot{Y}_D}{\partial \mathbf{x}_0} = \begin{bmatrix} \frac{2n\sqrt{1+s}(3+5s)}{1-s} \cos(nt\sqrt{1-s}) - \frac{2n\sqrt{1+s}(3+5s)}{1-s} + \mathbf{A}_3 \left\{ 1 + \frac{(3+5s)}{1-s} \right. \\ \left. - \frac{(3+5s)}{1-s} \cos(nt\sqrt{1-s}) \right\} \\ 0 \\ -\mathbf{A}_1 \left\{ \frac{\partial z}{\partial z_0} \right\} \\ \frac{-2\sqrt{1+s}}{\sqrt{1-s}} \sin(nt\sqrt{1-s}) + \mathbf{A}_3 \left\{ \frac{\sin(nt\sqrt{1-s})}{n\sqrt{1-s}} \right\} \\ \frac{4(1+s)}{1-s} \cos(nt\sqrt{1-s}) + 1 - \frac{4(1+s)}{1-s} + \mathbf{A}_3 \left\{ \frac{2\sqrt{1+s}}{n(1-s)} - \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) \right\} \\ -\mathbf{A}_1 \left\{ \frac{\partial z}{\partial z_0} \right\} \end{bmatrix}^T$$

$$\frac{\partial \dot{Z}_D}{\partial \mathbf{x}_0} = \begin{bmatrix} \mathbf{A}_1 \left\{ \frac{2\sqrt{1+s}(3+5s)}{\sqrt[3]{1-s}} \sin(nt\sqrt{1-s}) - \frac{2n\sqrt{1+s}(3+5s)}{1-s} t \right\} - \mathbf{A}_2 \left\{ 1 + \frac{(3+5s)}{1-s} - \frac{(3+5s)}{1-s} \right\} \\ \mathbf{A}_1 \left\{ \sin(qt + \phi) \frac{\partial l}{\partial z_0} + lt \cos(qt + \phi) \frac{\partial q}{\partial z_0} \right\} + \left\{ \frac{\partial \cos(qt + \phi)}{\partial z_0} + (lt + m) \cos(qt + \phi) \right\} \\ \left\{ \frac{\partial q}{\partial z_0} \right\} \\ \mathbf{A}_1 \left\{ \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) - \frac{2\sqrt{1+s}}{n(1-s)} \right\} - \mathbf{A}_2 \left\{ \frac{\sin(nt\sqrt{1-s})}{n\sqrt{1-s}} \right\} \\ \mathbf{A}_1 \left\{ \frac{4(1+s)}{n(\sqrt[3]{1-s})} \sin(nt\sqrt{1-s}) + t - \frac{4(1+s)}{1-s} t \right\} - \mathbf{A}_2 \left\{ \frac{2\sqrt{1+s}}{n(1-s)} - \frac{2\sqrt{1+s}}{n(1-s)} \cos(nt\sqrt{1-s}) \right\} \\ \left\{ \sin(qt + \phi) \frac{\partial l}{\partial \dot{z}_0} + l \cos(qt + \phi) \frac{\partial q}{\partial \dot{z}_0} \right\} + \left\{ \frac{\partial \cos(qt + \phi)}{\partial \dot{z}_0} + (lt + m) \cos(qt + \phi) \right\} \\ \left\{ \frac{\partial q}{\partial \dot{z}_0} \right\} \end{bmatrix}^T$$

where,

$$\frac{\partial l}{\partial z_0} = -r_C \sin I_D \sin I_C (\dot{\Omega}_D - \dot{\Omega}_C) \left(\frac{\cos \Delta \Omega_0 \sin \Phi_0 \frac{\partial \Delta \Omega_0}{\partial z_0} - \sin \Delta \Omega_0 \cos \Phi_0 \frac{\partial \Phi_0}{\partial z_0}}{\sin^2 \Phi_0} \right)$$

$$\frac{\partial \Phi_0}{\partial z_0} = \frac{1}{\sqrt{1 - (\cos I_D \cos I_C + \sin I_D \sin I_C \cos \Delta \Omega_0)^2}} (\sin I_D \sin I_C \sin \Delta \Omega_0 \frac{\partial \Delta \Omega_0}{\partial z_0})$$

$$\frac{\partial \Delta \Omega_0}{\partial z_0} = \frac{1}{r_C \sin I_C}, \quad \frac{\partial m}{\partial z_0} = \frac{1}{\sin \phi},$$

$$\frac{\partial \gamma_0}{\partial z_0} = -\frac{1}{\left(\frac{\cot I_C \sin I_D - \cos I_D \cos \Delta \Omega_0}{\sin \Delta \Omega_0}\right)^2 + 1}$$

$$\times \left(\frac{\cos I_D \sin^2 \Delta \Omega_0 \frac{\partial \Delta \Omega_0}{\partial z_0} - \cos \Delta \Omega_0 (\cot I_C \sin I_D - \cos I_D \cos \Delta \Omega_0) \frac{\partial \Delta \Omega_0}{\partial z_0}}{\sin^2 \Delta \Omega_0} \right)$$

$$\frac{\partial q}{\partial z_0} = \left\{ \begin{aligned} &(\dot{\Omega}_D - \dot{\Omega}_C)(-\sin^2 \gamma_0 \cot \Delta \Omega_0 \frac{\partial \gamma_0}{\partial z_0} + \cos^2 \gamma_0 \cot \Delta \Omega_0 \frac{\partial \gamma_0}{\partial z_0} - 2 \sin \gamma_0 \cos \gamma_0 \cos I_D \frac{\partial \gamma_0}{\partial z_0}) \\ &- \csc^2 \Delta \Omega_0 \cos \gamma_0 \sin \gamma_0 \frac{\partial \Delta \Omega_0}{\partial z_0} \end{aligned} \right\}$$

$$\frac{\partial z}{\partial z_0} = \left(\left(t \frac{\partial l}{\partial z_0} + \frac{\partial m}{\partial z_0} \right) \sin(qt + \phi) + t(lt + m) \cos(qt + \phi) \frac{\partial q}{\partial z_0} \right)$$

Now for Partial with respect to \dot{z}_0

$$\frac{\partial I_D}{\partial \dot{z}_0} = \frac{1}{kr_C}, \quad \frac{\partial \dot{\Omega}_D}{\partial \dot{z}_0} = \frac{3}{2kr_C} n J_2 \left(\frac{R}{r_C} \right)^2 \sin I_D$$

$$\frac{\partial \Phi_0}{\partial \dot{z}_0} = \frac{1}{\sqrt{1 - (\cos I_D \cos I_C + \sin I_D \sin I_C \cos \Delta \Omega_0)^2}}$$

$$\times (\cos I_C \sin I_D - \cos I_D \sin I_C \cos \Delta \Omega_0) \frac{\partial I_D}{\partial \dot{z}_0}$$

$$\frac{\partial l}{\partial \dot{z}_0} = -r_C \sin I_C \sin \Delta \Omega_0$$

$$\times \left\{ \frac{\cos I_D \sin \Phi_0 (\dot{\Omega}_D - \dot{\Omega}_C) \frac{\partial I_D}{\partial \dot{z}_0} + \cos \Phi_0 \sin I_D (\dot{\Omega}_C - \dot{\Omega}_D) \frac{\partial \Phi_0}{\partial \dot{z}_0} + \sin I_D \sin \Phi_0 \frac{\partial \dot{\Omega}_D}{\partial \dot{z}_0}}{\sin^2 \Phi_0} \right\}$$

$$\frac{\partial \gamma_0}{\partial \dot{z}_0} = -\frac{1}{\left(1 + \left(\frac{\cot I_C \sin I_D - \cos I_D \cos \Delta \Omega_0}{\sin \Delta \Omega_0} \right)^2 \right)} \left\{ \frac{\cot I_C \cos I_D \frac{\partial I_D}{\partial \dot{z}_0} + \sin I_D \cos \Delta \Omega_0 \frac{\partial I_D}{\partial \dot{z}_0}}{\sin \Delta \Omega_0} \right\}$$

Let suppose following:

$$u_1 = \cos \gamma_0 \sin \gamma_0 \cot \Delta \Omega_0 - \sin^2 \gamma_0 \cos I_D$$

$$\frac{\partial u_1}{\partial \dot{z}_0} = (-\sin^2 \gamma_0 \cot \Delta \Omega_0 + \cos^2 \gamma_0 \cot \Delta \Omega_0 - 2 \sin \gamma_0 \cos \gamma_0 \cos I_D) \times \frac{\partial \gamma_0}{\partial \dot{z}_0} + \sin I_D \sin^2 \gamma_0 \frac{\partial I_D}{\partial \dot{z}_0}$$

$$v_1 = \dot{\Omega}_D - \dot{\Omega}_C$$

$$\frac{\partial v_1}{\partial \dot{z}_0} = \frac{3}{2kr_C} nJ_2 \left(\frac{R}{r_C} \right)^2 \sin I_D$$

$$u_2 = \dot{\Omega}_D, \frac{\partial u_2}{\partial \dot{z}_0} = \frac{3}{2kr_C} nJ_2 \left(\frac{R}{r_C} \right)^2 \sin I_D$$

$$v_2 = \cos I_D, \frac{\partial v_2}{\partial \dot{z}_0} = -\sin I_D \frac{\partial I_D}{\partial \dot{z}_0}$$

Then,

$$\frac{\partial q}{\partial \dot{z}_0} = -(u'_1 v_1 + v'_1 u_1) - (u'_2 v_2 + v'_2 u_2)$$

$$\begin{aligned} \frac{\partial \dot{z}}{\partial \dot{z}_0} &= \{ \sin(qt + \phi) + qt \cos(qt + \phi) \} \frac{\partial l}{\partial \dot{z}_0} + \{ lt \cos(qt + \phi) - qt(lt + m) \sin(qt + \phi) \\ &+ (lt + m) \cos(qt + \phi) \} \frac{\partial q}{\partial \dot{z}_0} \end{aligned}$$

Appendix G: Analytical Solution of Integrals for GCSMM Time Update Equations

The analytical solutions presented in this Appendix are found out by making use of following important derivations:

- Product of two Gaussian densities $p_g^{(i)}$ and $p_g^{(j)}$:

$$\int_{-\infty}^{-\infty} \frac{\exp \left[-\frac{1}{2} \left((\mathbf{x} - \boldsymbol{\mu}^{(i)})^T \mathbf{P}^{-1(i)} (\mathbf{x} - \boldsymbol{\mu}^{(i)}) + (\mathbf{x} - \boldsymbol{\mu}^{(j)})^T \mathbf{P}^{-1(j)} (\mathbf{x} - \boldsymbol{\mu}^{(j)}) \right) \right]}{\sqrt{|2\pi(\mathbf{P}^{(i)} + \mathbf{P}^{(j)})|}} d\mathbf{x} \quad (\text{F.1})$$

The above integral can be re-expressed as:

$$\frac{\exp [K^{(i)}]}{(2\pi)^{\frac{d}{2}} \sqrt{|\mathbf{P}^{(i)} + \mathbf{P}^{(j)}|}} \int_{-\infty}^{-\infty} \exp \left[-\frac{1}{2} \left((\mathbf{e}^{(i)} - \mathbf{d}^{(i)})^T \boldsymbol{\Omega}^{-1} (\mathbf{e}^{(i)} - \mathbf{d}^{(i)}) \right) \right] d\mathbf{e}^{(i)} \quad (\text{F.2})$$

where,

$$\begin{aligned} K^{(i)} &= \frac{1}{2} \mathbf{z}^{(i)T} (\mathbf{P}^{-1(j)} \boldsymbol{\Omega} \mathbf{P}^{-1(j)} - \mathbf{P}^{-1(j)}) \mathbf{z}^{(i)} \\ \mathbf{e}^{(i)} &= \mathbf{x} - \boldsymbol{\mu}^{(i)} \\ \mathbf{z}^{(i)} &= \boldsymbol{\mu}^{(j)} - \boldsymbol{\mu}^{(i)} \\ \mathbf{d}^{(i)} &= \boldsymbol{\Omega} \mathbf{P}^{-1(j)} \mathbf{z}^{(i)} \\ \boldsymbol{\Omega} &= (\mathbf{P}^{-1(i)} + \mathbf{P}^{-1(j)})^{-1} \end{aligned}$$

- Results for Gaussian based expectation integrals [1]:

$$I_0 = \int_{-\infty}^{-\infty} \exp \left[-\frac{1}{2} \left((\mathbf{x}_r - \mathbf{b}_r)^T \boldsymbol{\Omega}_{rs}^{-1} (\mathbf{x}_s - \mathbf{b}_s) \right) \right] d\mathbf{x} = (2\pi)^{\frac{d}{2}} |\boldsymbol{\Omega}_{ij}|^{1/2} \quad (\text{F.3})$$

$$\begin{aligned}
I_1 &= \int_{-\infty}^{-\infty} \exp \left[-\frac{1}{2} \left((\mathbf{x}_r - \mathbf{b}_r)^T \Omega_{rs}^{-1} (\mathbf{x}_s - \mathbf{b}_s) \right) \right] (\mathbf{x}_i - \mathbf{b}_i) d\mathbf{x} = 0 \\
I_2 &= \int_{-\infty}^{-\infty} \exp \left[-\frac{1}{2} \left((\mathbf{x}_r - \mathbf{b}_r)^T \Omega_{rs}^{-1} (\mathbf{x}_s - \mathbf{b}_s) \right) \right] (\mathbf{x}_i - \mathbf{b}_i) (\mathbf{x}_j - \mathbf{b}_j) d\mathbf{x} \\
&= (2\pi)^{\frac{d}{2}} |\Omega_{ij}|^{1/2} \Omega_{ij} \\
I_3 &= \int_{-\infty}^{-\infty} \exp \left[-\frac{1}{2} \left((\mathbf{x}_r - \mathbf{b}_r)^T \Omega_{rs}^{-1} (\mathbf{x}_s - \mathbf{b}_s) \right) \right] (\mathbf{x}_i - \mathbf{b}_i) (\mathbf{x}_j - \mathbf{b}_j) (\mathbf{x}_k - \mathbf{b}_k) d\mathbf{x} = 0 \\
I_4 &= \int_{-\infty}^{-\infty} \exp \left[-\frac{1}{2} \left((\mathbf{x}_r - \mathbf{b}_r)^T \Omega_{rs}^{-1} (\mathbf{x}_s - \mathbf{b}_s) \right) \right] (\mathbf{x}_i - \mathbf{b}_i) (\mathbf{x}_j - \mathbf{b}_j) (\mathbf{x}_k - \mathbf{b}_k) (\mathbf{x}_l - \mathbf{b}_l) d\mathbf{x} \\
&= (2\pi)^{\frac{d}{2}} |\Omega_{ij}|^{1/2} (\Omega_{ij} \Omega_{kl} + \Omega_{ik} \Omega_{jl} + \Omega_{il} \Omega_{jk})
\end{aligned}$$

where, $|\Omega_{ij}|^{1/2}$ = determinant of $\mathbf{\Omega}$

- The components of matrix \mathbf{M}_c (Equation: 6.44)

$$m_{cij} = \frac{1}{\Delta t^2} \left[\frac{\exp [K^{(j)}]}{(2\pi)^{\frac{d}{2}} \sqrt{|\mathbf{P}^{(i)} + \mathbf{P}^{(j)}|}} + \frac{\exp [K^{(j)}]}{\sqrt{|\mathbf{P}^{(i)} + \mathbf{P}^{(j)}|}} |\Omega_{ij}|^{1/2} \mathcal{V}^{(j)} + \frac{\exp [K^{(i)}]}{\sqrt{|\mathbf{P}^{(i)} + \mathbf{P}^{(j)}|}} |\Omega_{ij}|^{1/2} \mathcal{V}^{(i)} \right] \quad (\text{F.4})$$

where, the term for j^{th} component are shown below:

$$K^{(j)} = \frac{1}{2} \mathbf{z}^{(j)T} (\mathbf{P}^{-1(i)} \mathbf{\Omega} \mathbf{P}^{-1(i)} - \mathbf{P}^{-1(i)}) \mathbf{z}^{(j)}$$

$$\mathbf{e}^{(j)} = \mathbf{x} - \boldsymbol{\mu}^{(j)}$$

$$\mathbf{z}^{(j)} = \boldsymbol{\mu}^{(i)} - \boldsymbol{\mu}^{(j)}$$

$$\mathbf{d}^{(j)} = \mathbf{\Omega} \mathbf{P}^{-1(i)} \mathbf{z}^{(j)}$$

$$\mathbf{\Omega} = (\mathbf{P}^{-1(i)} + \mathbf{P}^{-1(j)})^{-1}$$

$$\mathcal{V}^{(j)} = \frac{1}{2} \mathbf{P}_{abc}^{(3)(j)} \eta_a^{(j)} \mathbf{E}_{bc}^{(j)}$$

$$\eta_a^{(j)} = \mathbf{P}_{ar}^{-1(j)} \mathbf{d}_r^{(j)}$$

$$\mathbf{G}_{ab}^{(j)} = \Omega_{ae} \mathbf{P}_{eb}^{-1(j)}$$

$$\mathbf{Q}_{ab}^{(j)} = \mathbf{P}_{ae}^{-1(j)} \mathbf{G}_{eb}^{(j)}$$

$$\mathbf{E}_{bc}^{(j)} = \mathbf{Q}_{ab}^{(j)} - \mathbf{P}_{ab}^{-1(j)} + \frac{1}{3} \eta_a^{(j)} \eta_b^{(j)}$$

(F.4a)

For $K^{(i)}$ and $\mathcal{V}^{(i)}$ replace “ j ” by “ i ” in Equation: F.4a.

For $i = j$:

$$m_{c_{ii}} = \frac{1}{\Delta t^2} \left[\frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{|2\mathbf{P}^{(i)}|}} \right] \quad (\text{F.5})$$

- The components of matrix \mathbf{N}_c are expressed as:

$$\begin{aligned} n_{c_{ij}} = & \frac{1}{\Delta t} \int_{-\infty}^{-\infty} p_{gcs}^{(i)} \left[\left(\frac{\partial p_{gcs}^{(j)T}}{\partial \boldsymbol{\mu}^{(j)}} \dot{\boldsymbol{\mu}}^{(j)} \right) + \text{Tr} \left[\frac{\partial p_{gcs}^{(j)T}}{\partial \mathbf{P}^{(j)}} \dot{\mathbf{P}}^{(j)} \right] + \frac{\partial p_{gcs}^{(j)}}{\partial P_{abc}^{(3)(j)}} \dot{P}_{abc}^{(3)(j)} - \frac{1}{\Delta t} p_{gcs}^{(j)} + \frac{\partial p_{gcs}^{(j)T}}{\partial \mathbf{x}} \mathbf{f}(t, \mathbf{x}) \right. \\ & \left. + p_{gcs}^{(j)} \text{Tr} \left[\frac{\partial \mathbf{f}(t, \mathbf{x})}{\partial \mathbf{x}} \right] - \frac{1}{2} \text{Tr} \left[\mathbb{Q} \frac{\partial^2 p_{gcs}^{(j)}}{\partial \mathbf{x} \partial \mathbf{x}^T} \right] \right] d\mathbf{x} \end{aligned} \quad (\text{F.6})$$

We shall utilize tensor notation to solve above integral analytically. Each of the above term inside the square bracket of integrand can be treated separately:

$$n_{c_{ij}}^1 = \frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \left[\left(\frac{\partial p_{gcs}^{(j)T}}{\partial \boldsymbol{\mu}^j} \dot{\boldsymbol{\mu}}^j \right) \right] d\mathbf{x} \quad (\text{F.7})$$

Substituting Equation: 6.31 and taking expectation of the function inside square bracket (Equation: F.7) and making use of results given in Equations: F.1-F.3 we obtain following:

$$\begin{aligned}
n_{c_{ij}}^1 = & \frac{1}{\Delta t} \frac{|\Omega_{ij}|^{1/2} \dot{\mu}_a^{(j)}}{\sqrt{|\mathbf{P}^{(i)} + \mathbf{P}^{(j)}|}} \left\{ \exp [\mathbf{K}^{(j)}] \left(\eta_a^{(j)} \right. \right. \\
& + \frac{\mathbf{P}_{lmn}^{(3)(j)}}{3!} \left[\mathbf{Q}_{lm}^{(j)} \mathbf{Q}_{na}^{(j)} [3] + \eta_l^{(j)} \eta_m^{(j)} \eta_n^{(j)} \eta_a^{(j)} + \mathbf{Q}_{lm}^{(j)} \eta_n^{(j)} \eta_a^{(j)} [3] + \mathbf{Q}_{la}^{(j)} \eta_m^{(j)} \eta_n^{(j)} [3] \right. \\
& - \mathbf{P}_{mn}^{-1(j)} \mathbf{Q}_{la}^{(j)} [3] - \mathbf{P}_{mn}^{-1(j)} \eta_l^{(j)} \eta_a^{(j)} [3] - \mathbf{P}_{la}^{-1(j)} \mathbf{Q}_{mn}^{(j)} [3] - \mathbf{P}_{la}^{-1(j)} \eta_m^{(j)} \eta_n^{(j)} [3] \\
& \left. \left. + \mathbf{P}_{la}^{-1(j)} \mathbf{P}_{mn}^{-1(j)} [3] \right] \right) \\
& + \exp [\mathbf{K}^{(i)}] \frac{\mathbf{P}_{lmn}^{(3)(i)}}{3!} \left[\mathbf{Q}_{lm}^{(i)} \mathbf{Q}_{na}^{(ij)} [3] + \eta_l^{(i)} \eta_m^{(i)} \eta_n^{(i)} \zeta_a^{(ji)} + \mathbf{Q}_{lm}^{(i)} \eta_n^{(i)} \zeta_a^{(ji)} [3] \right. \\
& \left. \left. + \mathbf{Q}_{na}^{(ij)} \eta_l^{(i)} \eta_m^{(i)} [3] - \mathbf{P}_{mn}^{-1(i)} \mathbf{Q}_{la}^{(ij)} [3] - \mathbf{P}_{mn}^{-1(i)} \eta_l^{(i)} \zeta_a^{(ji)} [3] \right] \right\}
\end{aligned}$$

where,

double superscript variables are:

$$\begin{aligned}
\mathbf{Q}_{na}^{(ij)} & \in \mathbf{P}^{-1(i)} \mathbf{\Omega} \mathbf{P}^{-1(j)} \\
\zeta_a^{(ji)} & = \mathbf{P}_{au}^{-1(j)} \left(\mathbf{d}_u^{(i)} - \mathbf{z}_u^{(i)} \right) \\
\eta_a^{(ji)} & = \mathbf{P}_{au}^{-1(j)} \mathbf{d}_u^{(i)}
\end{aligned}$$

Other variables used in above expression are similar to Equation: F.4. Now we solve the second integrand as:

$$n_{c_{ij}}^2 = \frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \text{Tr} \left[\frac{\partial p_{gcs}^{(j)T}}{\partial \mathbf{P}^{(j)}} \dot{\mathbf{P}}^{(j)} \right] d\mathbf{x} \tag{F.8}$$

By substituting Equation: 6.32 in above equation the solution can be expressed as:

$$\begin{aligned}
n_{c_{ij}}^2 = & \frac{1}{\Delta t} \frac{|\Omega_{ij}|^{1/2} \dot{P}_{ml}^{(j)}}{\sqrt{|(\mathbf{P}^{(i)} + \mathbf{P}^{(j)})|}} \left\{ \exp[K^{(j)}] \left(\frac{1}{2} [Q_{lm}^{(j)} + \eta_l^{(j)} \eta_m^{(j)} - P_{lm}^{-1(j)}] \right. \right. \\
& - \frac{P_{abc}^{(3)(j)}}{3!} \left[\frac{1}{2} \left((\eta_l^{(j)} Q_{ma}^{(j)} [3]) P_{bc}^{-1(j)} + \eta_l^{(j)} \eta_m^{(j)} \eta_a^{(j)} P_{bc}^{-1(j)} + (\eta_l^{(j)} Q_{mb}^{(j)} [3]) P_{ac}^{-1(j)} \right. \right. \\
& + \eta_l^{(j)} \eta_m^{(j)} \eta_b^{(j)} P_{ac}^{-1(j)} + (\eta_l^{(j)} Q_{mc}^{(j)} [3]) P_{ab}^{-1(j)} + \eta_l^{(j)} \eta_m^{(j)} \eta_c^{(j)} P_{ab}^{-1(j)} \\
& + (\eta_a^{(j)} Q_{bc}^{(j)} [3]) P_{lm}^{-1(j)} + \eta_a^{(j)} \eta_b^{(j)} \eta_c^{(j)} P_{lm}^{-1(j)} \Big) - \frac{1}{2} (P_{lm}^{-1(j)} (P_{ab}^{-1(j)} \eta_c^{(j)} [3])) \\
& + (\eta_m^{(j)} Q_{bc}^{(j)} [3]) P_{al}^{-1(j)} + \eta_m^{(j)} \eta_b^{(j)} \eta_c^{(j)} P_{al}^{-1(j)} + (\eta_a^{(j)} Q_{mc}^{(j)} [3]) P_{bl}^{-1(j)} \\
& + \eta_a^{(j)} \eta_m^{(j)} \eta_c^{(j)} P_{bl}^{-1(j)} + (\eta_a^{(j)} Q_{mb}^{(j)} [3]) P_{cl}^{-1(j)} + \eta_a^{(j)} \eta_b^{(j)} \eta_m^{(j)} P_{cl}^{-1(j)} \\
& \left. \left. - (P_{al}^{-1(j)} P_{bc}^{-1(j)} [3]) \eta_m^{(j)} - (P_{bl}^{-1(j)} P_{mc}^{-1(j)} \eta_a^{(j)} [3]) \right] \right) \\
& - \frac{\exp[K^{(i)}] P_{abc}^{(3)(i)}}{3!} \left(\frac{1}{2} \left((\eta_a^{(i)} P_{bc}^{-1(i)} [3]) Q_{lm}^{(j)} + (\eta_a^{(i)} P_{bc}^{-1(i)} [3]) \zeta_l^{(ji)} \zeta_m^{(ji)} \right. \right. \\
& + ((P_{ab}^{-1(i)} Q_{mc}^{(ij)} [3]) \zeta_l^{(ji)}) + (P_{ab}^{-1(i)} Q_{lc}^{(ji)} [3]) \zeta_m^{(ji)} + P_{lm}^{-1(j)} ((\eta_a^{(i)} Q_{bc}^{(i)} [3]) \\
& \left. \left. + \eta_a^{(i)} \eta_b^{(i)} \eta_c^{(i)} - \eta_a^{(i)} P_{bc}^{-1(i)} [3]) \right) \right) \Big\}
\end{aligned}$$

where,

$$Q_{mc}^{(ji)} \in \mathbf{P}^{-1(j)} \mathbf{\Omega} \mathbf{P}^{-1(i)}$$

Now we solve the third integrand inside square bracket of Equation: F.6. The integrand can be written as:

$$n_{c_{ij}}^3 = \frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \frac{\partial p_{gcs}^{(j)}}{\partial P_{abc}^{(3)(j)}} \dot{P}_{abc}^{(3)(j)} d\mathbf{x} \quad (\text{F.9})$$

Substituting Equation: 6.35 and using the results of Equation: F.1 and F.2 we perform expectation of above integral with respect $p_{gcs}^{(i)}$. The solution can be expressed as:

$$n_{c_{ij}}^3 = \frac{1}{\Delta t} \frac{|\Omega_{ij}|^{1/2} \exp[K^{(j)}]}{3! \sqrt{|(\mathbf{P}^{(i)} + \mathbf{P}^{(j)})|}} \left[\dot{P}_{abc}^{(3)(j)} (\eta_a^{(j)} E_{bc}^{(j)} [3]) \right] \quad (\text{F.10})$$

See Equation: F.4 for solution of following integral as these are identical:

$$n_{cij}^4 = \frac{1}{\Delta t^2} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} p_{gcs}^{(j)} d\mathbf{x}$$

Now we solve for following integral:

$$n_{cij}^5 = \frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \frac{\partial p_{gcs}^{(j)T}}{\partial \mathbf{x}} \mathbf{f}(t, \mathbf{x}) d\mathbf{x} \quad (\text{F.11})$$

The solution of the above integral can be simplified by expanding the nonlinear function up to second order in Taylor series and substituting Equation: 6.33 in F.11. The solution can be written as:

$$\begin{aligned} n_{cij}^5 = \frac{1}{\Delta t} \frac{|\Omega_{ij}|^{1/2}}{\sqrt{|(\mathbf{P}^{(i)} + \mathbf{P}^{(j)})|}} & \left[\exp[K^{(j)}] \left(-P_{au}^{-1(j)} \left(d_u^{(j)} f_a(t, \boldsymbol{\mu}^{(j)}) + F_{aw}^{(j)} (\Omega_{uw} + d_u^{(j)} d_w^{(j)}) \right. \right. \right. \\ & + A_{awf}^{(j)} (d_u^{(j)} \Omega_{wf} + d_w^{(j)} \Omega_{uf} + d_f^{(j)} \Omega_{uw} + d_u^{(j)} d_w^{(j)} d_f^{(j)}) \\ & - \frac{P_{lmn}^{(3)(j)}}{3!} \left(f_a(t, \boldsymbol{\mu}^{(j)}) \left(Q_{lm}^{(j)} Q_{na}^{(j)} [3] + \eta_l^{(j)} \eta_m^{(j)} \eta_n^{(j)} \eta_a^{(j)} + Q_{lm}^{(j)} \eta_n^{(j)} \eta_a^{(j)} [3] \right. \right. \\ & + Q_{la}^{(j)} \eta_m^{(j)} \eta_n^{(j)} [3] - P_{mn}^{-1(j)} Q_{la}^{(j)} [3] - P_{mn}^{-1(j)} \eta_l^{(j)} \eta_a^{(j)} [3] \\ & - \left(P_{la}^{-1(j)} Q_{mn}^{(j)} + P_{la}^{-1(j)} \eta_m^{(j)} \eta_n^{(j)} \right) [3] \Big) \\ & - \left(P_{la}^{-1(j)} \left(\eta_m^{(j)} L_{na}^{(j)} + \eta_n^{(j)} L_{ma}^{(j)} + \xi_a^{(j)} Q_{mn}^{(j)} + \eta_m^{(j)} \eta_n^{(j)} \xi_a^{(j)} \right) \right) [3] \\ & + \left(f_a(t, \boldsymbol{\mu}^{(j)}) + F_{aw}^{(j)} d_w^{(j)} + A_{awf}^{(j)} (\Omega_{wf} + d_w^{(j)} d_f^{(j)}) \right) \left(P_{la}^{-1(j)} P_{mn}^{-1(j)} [3] \right) \Big) \\ & + \frac{\exp[K^{(i)}] P_{lmn}^{(3)(i)}}{3!} \left(f_a(t, \boldsymbol{\mu}^{(j)}) (P_{lm}^{-1(i)} Q_{an}^{(ij)} [3] + \eta_l^{(i)} \zeta_a^{(ji)} P_{mn}^{-1(i)} [3]) \right. \\ & \left. + \left(L_{aa}^{(j)} \eta_l^{(i)} P_{mn}^{-1(i)} + Q_{al}^{(ji)} P_{mn}^{-1(i)} \xi_a^{(ji)} + \zeta_a^{(ji)} L_{al}^{(ij)} P_{mn}^{-1(i)} + \eta_l^{(i)} P_{mn}^{-1(i)} \xi_a^{(ji)} \zeta_a^{(ji)} \right) [3] \right) \Big] \end{aligned}$$

where, the new variables defined in above equation are:

$$F_{aw}^{(j)} = \frac{\partial f_a(t, \mathbf{x})}{\partial x_w} \Big|_{\mathbf{x}=\boldsymbol{\mu}^{(j)}}, \quad A_{awf}^{(j)} = \frac{1}{2} \frac{\partial^2 f_a(t, \mathbf{x})}{\partial x_w \partial x_f} \Big|_{\mathbf{x}=\boldsymbol{\mu}^{(j)}}$$

$$\mathbf{L}^{(j)} = \mathbf{P}^{(j)} \mathbf{\Omega} \mathbf{F}^{T(j)}, \quad \mathbf{L}^{(ij)} = \mathbf{P}^{(i)} \mathbf{\Omega} \mathbf{F}^{T(j)}$$

$$\xi_a^{(j)} = F_{aw}^{(j)} d_w^{(j)}, \quad \xi_a^{(ji)} = F_{aw}^{(j)} (d_w^{(i)} - z_w^{(i)})$$

$$n_{c_{ij}}^6 = \frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} p_{gcs}^{(j)} \text{Tr} \left[\frac{\partial \mathbf{f}(t, \mathbf{x})}{\partial \mathbf{x}} \right] d\mathbf{x} \quad (\text{F.12})$$

The integral in Equation: F.12 can also be solved on similar lines as previously described methodology of Equation: F.11. However, in our orbit determination application (Section: 6.5) this integral is zero. Therefore, it will not be treated further.

Now we solve for following integral:

$$n_{c_{ij}}^6 = -\frac{1}{2\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \frac{1}{2} \text{Tr} \left[\mathbb{Q} \frac{\partial^2 p_{gcs}^{(j)}}{\partial \mathbf{x} \partial \mathbf{x}^T} \right] d\mathbf{x} \quad (\text{F.13})$$

By substituting Equation: 6.34 in Equation: F.13, the solution of above integral can be written as:

$$\begin{aligned} n_{c_{ij}}^6 = & -\frac{1}{2\Delta t} \frac{|\Omega_{ij}|^{1/2} \mathbb{Q}_{ml}}{\sqrt{|\mathbf{P}^{(i)} + \mathbf{P}^{(j)}|}} \left[\exp[\mathbf{K}^{(j)}] \left(\left(\mathbf{Q}_{lm}^{(j)} + \eta_l^{(j)} \eta_m^{(j)} - \mathbf{P}_{lm}^{-1(j)} \right) \right. \right. \\ & - \frac{\mathbf{P}_{abc}^{(3)(j)}}{3!} \left((\eta_l^{(j)} \mathbf{Q}_{ma}^{(j)} [3]) \mathbf{P}_{bc}^{-1(j)} + (\eta_l^{(j)} \mathbf{Q}_{mb}^{(j)} [3]) \mathbf{P}_{ac}^{-1(j)} + (\eta_l^{(j)} \mathbf{Q}_{mc}^{(j)} [3]) \mathbf{P}_{ab}^{-1(j)} \right. \\ & + (\eta_a^{(j)} \mathbf{Q}_{bc}^{(j)} [3]) \mathbf{P}_{lm}^{-1(j)} + \eta_l^{(j)} \eta_m^{(j)} (\eta_a^{(j)} \mathbf{P}_{bc}^{-1(j)} [3]) + \eta_a^{(j)} \eta_b^{(j)} \eta_c^{(j)} \mathbf{P}_{lm}^{-1(j)} \\ & \left. \left. - \mathbf{P}_{lm}^{-1(j)} (\eta_a^{(j)} \mathbf{P}_{bc}^{-1(j)} [3]) - \mathbf{P}_{al}^{-1(j)} \mathbf{P}_{bm}^{-1(j)} \eta_c^{(j)} [6] \right) \right) \\ & - \frac{\exp[\mathbf{K}^{(i)}] \mathbf{P}_{abc}^{(3)(i)}}{3!} \left(\left((\eta_a^{(i)} \mathbf{P}_{bc}^{-1(i)} [3]) \mathbf{Q}_{lm}^{(j)} + (\eta_a^{(i)} \mathbf{P}_{bc}^{-1(i)} [3]) \zeta_l^{(ji)} \zeta_m^{(ji)} \right. \right. \\ & + (\mathbf{P}_{ab}^{-1(i)} \mathbf{Q}_{mc}^{(ij)} [3]) \zeta_l^{(ji)} \left. \left. + (\mathbf{P}_{ab}^{-1(i)} \mathbf{Q}_{lc}^{(ij)} [3]) \zeta_m^{(ji)} + \mathbf{P}_{lm}^{-1(j)} ((\eta_a^{(i)} \mathbf{Q}_{bc}^{(i)} [3]) \right. \right. \\ & \left. \left. + \eta_a^{(i)} \eta_b^{(i)} \eta_c^{(i)} - \eta_a^{(i)} \mathbf{P}_{bc}^{-1(i)} [3]) \right) \right) \right] \quad (\text{F.14}) \end{aligned}$$

Now for $i=j$ the components of matrix \mathbf{N}_c are expressed as:

$$n_{cii} = \frac{1}{\Delta t} \int_{-\infty}^{-\infty} p_{gcs}^{(i)} \left[\left(\frac{\partial p_{gcs}^{(i)T}}{\partial \mathbf{\mu}^{(i)}} \dot{\mathbf{\mu}}^{(i)} \right) + \text{Tr} \left[\frac{\partial p_{gcs}^{(i)T}}{\partial \mathbf{P}^{(i)}} \dot{\mathbf{P}}^{(i)} \right] + \frac{\partial p_{gcs}^{(i)}}{\partial P_{abc}^{(3)(i)}} \dot{P}_{abc}^{(3)(i)} - \frac{1}{\Delta t} p_{gcs}^{(i)} + \frac{\partial p_{gcs}^{(i)T}}{\partial \mathbf{x}} \mathbf{f}(t, \mathbf{x}) \right. \\ \left. + p_{gcs}^{(i)} \text{Tr} \left[\frac{\partial \mathbf{f}(t, \mathbf{x})}{\partial \mathbf{x}} \right] - \frac{1}{2} \text{Tr} \left[\mathbb{Q} \frac{\partial^2 p_{gcs}^{(i)}}{\partial \mathbf{x} \partial \mathbf{x}^T} \right] \right] d\mathbf{x} \quad (\text{F.15})$$

We shall utilize tensor notation to solve above integral analytically. Each of the above term inside the square bracket of integrand can be treated separately:

$$n_{cii}^1 = \frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \left[\left(\frac{\partial p_{gcs}^{(i)T}}{\partial \mathbf{\mu}^i} \dot{\mathbf{\mu}}^i \right) \right] d\mathbf{x} \quad (\text{F.16})$$

Substituting Equation: 6.31 and taking expectation of the function inside square bracket (Equation: F.16) and making use of results given in Equations: F.1-F.3 we obtain following:

$$n_{cii}^1 = -\frac{1}{\Delta t} \frac{|\mathbf{P}^{(i)}|^{1/2} \dot{\mu}_a^{(i)}}{2^{(d/2)} \sqrt{|(2\mathbf{P}^{(i)})|}} \left\{ \left(\frac{P_{lmn}^{(3)(i)}}{3!} \left[2P_{lm}^{(i)} Q_{an}^{(i)}[3] + P_{la}^{(i)} Q_{mn}^{(i)}[3] - P_{la}^{(i)} P_{mn}^{(i)}[3] \right] \right) \right\}$$

Now we solve the second integrand as:

$$n_{cii}^2 = \frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \text{Tr} \left[\frac{\partial p_{gcs}^{(i)T}}{\partial \mathbf{P}^{(i)}} \dot{\mathbf{P}}^{(i)} \right] d\mathbf{x} \quad (\text{F.17})$$

By substituting Equation: 6.32 in above equation the solution can be expressed as:

$$n_{cii}^2 = \frac{1}{\Delta t} \frac{|\mathbf{P}^{(i)}|^{1/2} \dot{P}_{ml}^{(i)}}{2^{(d/2)} \sqrt{|(2\mathbf{P}^{(i)})|}} \left\{ \left(\frac{1}{2} [Q_{lm}^{(i)} - P_{lm}^{-1(i)}] \right) \right\}$$

Now we solve the third integrand inside square bracket of Equation: F.15. The integrand can be written as:

$$n_{c_{ii}}^3 = \frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \frac{\partial p_{gcs}^{(i)}}{\partial \mathbf{P}_{abc}^{(3)(i)}} \dot{\mathbf{P}}_{abc}^{(3)(i)} d\mathbf{x} \quad (\text{F.18})$$

Substituting Equation: 6.35 and using the results of Equation: F.1 and F.2 we perform expectation of above integral with respect $p_{gcs}^{(i)}$. The solution can be expressed as:

$$n_{c_{ii}}^3 = 0 \quad (\text{F.19})$$

The fourth term inside square bracket of Equation: F.15 can be expressed as:

$$n_{c_{ii}}^4 = \frac{1}{\Delta t^2} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} p_{gcs}^{(i)} d\mathbf{x} = \frac{1}{\Delta t^2} \frac{|\mathbf{P}^{(i)}|^{1/2}}{(2)^{\frac{d}{2}} \sqrt{|2\mathbf{P}^{(i)}|}}$$

Now we solve for following integral:

$$n_{c_{ii}}^5 = \frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \frac{\partial p_{gcs}^{(i)T}}{\partial \mathbf{x}} \mathbf{f}(t, \mathbf{x}) d\mathbf{x} \quad (\text{F.20})$$

The solution of the above integral can be simplified by expanding the nonlinear function up to second order in Taylor series and substituting Equation: 6.33 in F.20. The solution can be written as:

$$n_{c_{ii}}^5 = \frac{1}{\Delta t} \frac{|\mathbf{P}^{(i)}|^{1/2}}{(2)^{\frac{d}{2}} \sqrt{|(2\mathbf{P}^{(i)})|}} \left[\left(-\mathbf{P}_{au}^{-1(i)} \mathbf{F}_{aw}^{(i)} \mathbf{P}_{uw}^{-1(i)} \right. \right. \\ \left. \left. + \frac{\mathbf{P}_{lmn}^{(3)(i)}}{3!} \left(\mathbf{f}_a(t, \boldsymbol{\mu}^{(i)}) \left(\mathbf{P}_{lm}^{(i)} \mathbf{Q}_{an}^{(i)}[3] + \mathbf{P}_{na}^{(i)} \mathbf{Q}_{lm}^{(i)}[3] \right) \right. \right. \right. \\ \left. \left. \left. - \left(\mathbf{f}_a(t, \boldsymbol{\mu}^{(i)}) + \mathbf{A}_{awf}^{(i)} \mathbf{P}_{wf}^{-1(i)} \right) \left(\mathbf{P}_{la}^{-1(i)} \mathbf{P}_{mn}^{-1(i)}[3] \right) \right) \right) \right]$$

Now we solve for following integral:

$$n_{c_{ii}}^6 = -\frac{1}{\Delta t} \int_{-\infty}^{+\infty} p_{gcs}^{(i)} \frac{1}{2} \text{Tr} \left[\mathbb{Q} \frac{\partial^2 p_{gcs}^{(i)}}{\partial \mathbf{x} \partial \mathbf{x}^T} \right] d\mathbf{x} \quad (\text{F.21})$$

By substituting Equation: 6.34 in Equation: F.13, the solution of above integral can be written as:

$$n_{c_{ii}}^6 = -\frac{1}{2\Delta t} \frac{|\mathbf{P}^{(i)}|^{1/2} \mathbb{Q}_{ml}}{(2)^{\frac{d}{2}} \sqrt{|(2\mathbf{P}^{(i)})|}} \left[\mathbb{Q}_{lm}^{(i)} - \mathbb{P}_{lm}^{-1(i)} \right]$$

In solving above integrals fourth and higher order moments and multiplicative terms involving their differentials are neglected.