Bayesian Estimation For Tracking Of Spiraling Reentry Vehicles

Juan Esteban Tapiero Bernal
Marquette University

BAYESIAN ESTIMATION FOR TRACKING OF SPIRALING REENTRY VEHICLES

by

Juan E. Tapiero Bernal, B.S.

A Thesis Submitted to the Faculty of the Graduate School
Marquette University,
in Partial Fulfillment of the Requirements for
the Degree of Master of Science

Milwaukee, Wisconsin

May 2013

ABSTRACT

BAYESIAN ESTIMATION FOR TRACKING OF
SPIRALING REENTRY VEHICLES

Juan E. Tapiero Bernal, B.S.
Marquette University

This thesis presents a development of a physics-based dynamics model of a spiraling atmospheric reentry vehicle. An analysis of the trajectory characteristics, using elements from differential geometry lead to a relationship of the state of the vehicle to the spiraling of motion. The Bayesian estimation framework for nonlinear systems is introduced showing the theoretical basis of the estimation techniques. Two estimation algorithms, extended Kalman filter and particle filter are presented, their mathematical formulation and implementation characteristics.

Different trajectories that can be represented by the model are introduced and analyzed, showing the spiraling behavior that can be described by the model. The extended Kalman filter and particle filter are compared in the ability to estimate the states and spiraling characteristics, with successful results for both techniques inside one standard deviation. In general, superior performance was shown by the particle filter which estimated the torsion with an error 10 orders of magnitude smaller.

ACKNOWLEDGEMENTS

I wish to express my gratitude to my advisor Dr. Robert H. Bishop, who has trusted me and provided me with academic and economic support for this research and my graduate studies. I would also like to thank Dr. Cesar Ocampo from The University of Texas at Austin and Jorge Alvarez who first supported me and facilitated my opportunity to become a graduate student in the United States, and all my friends at The University of Texas at Austin and Marquette University.

I would also like to thank the people that supported this work and helped exchanging ideas and knowledge about estimation techniques, in particular Adam Mallen who gave me some insight with the particle filter. Special thanks also to my committee members that gave me feedback to get a solid final thesis document. Finally I would like to thank my family that was and is always there for me.

TABLE OF CONTENTS

TABLE OF CONTENTS — *Continued*

TABLE OF CONTENTS — *Continued*

LIST OF TABLES

LIST OF FIGURES

LIST OF NOMENCLATURE AND ACRONYMS

PCI: Planet-Centered Inertial reference Frame.

KF: Kalman Filter.

EKF: Extended Kalman Filter.

CDEKF: Continuous-Discrete Extended Kalman Filter.

PF: Particle Filter.

CDPF: Continuous-Discrete Particle Filter.

SDE: Stochastic Differential Equation.

Table 0.1: Nomenclature

math	definition
boldface	vector or Matrix
italic	vector magnitude or scalar
lowercase	vector
UPPERCASE	Matrix
$\odot$	dot product
$*$	cross product
$\| \bullet \|$	standard Euclidean norm of a vector
β_m	ballistic coefficient
C_L	lift coefficient
C_D	drag coefficient
S	Effective area of the vehicle
m	Mass (Chapter 2) and in general dimension of observation vector
β_m	Ballistic coefficient
$p(\bullet)$	Probability density function
$N(\bullet)$	Normal Distribution
$\lvert \bullet \rvert$	Determinant of a matrix
n	number of states
N	number of particles or samples

CHAPTER 1

Problem Statement, Objective and Contributions

1.1 Problem statement

This thesis develops a mathematical model of an atmospheric reentry vehicle with coarse dynamics based on physical principles that can be successfully used to estimate position, velocity, acceleration, as well as the spiraling frequency of the vehicle. Key parameters, such as the ballistic coefficient, maximum lift-to-drag coefficient, ratio of lift to critical lift, and bank angle are also estimated. Achieving this through the use of Bayesian estimation techniques is a challenge given the nature of the significant nonlinearities present in the system.

This work has application in missile tracking and in the recovery of reentry unmanned vehicles from space missions. It is also a contribution that can be extrapolated to other application of tracking of moving objects, especially the use of Bayesian techniques in the state estimation of systems derived through physical principles. The Bayesian estimation techniques employed in this thesis are the extended Kalman filter and the particle filter.

1.2 Objectives

The main objective of this work is to show that the proposed mathematical model can be used for estimation of the state of a reentry vehicle, even though it possesses high

dimensionality and nonlinearity. The principles under which the model is established are based on physical principle and an excellent approximation to the real system motion. The second objective is to show the power of the Bayesian framework in creating practical estimation algorithms. In particular, the application of Bayesian estimation using a physics-based model to represent the dynamics is proved to be an outstanding approach using realistic simulation-based analysis.

1.3 Previous Work

Trajectories of reentry vehicles have been studied since the early days of the space program and the development of ballistic missiles (e.g. see [1];[2]). However estimation of the spiraling component of the trajectory of an endoatmospheric reentry vehicle is not a widely studied problem. The earliest known work on the estimation of the spiraling components of a vehicle are related to missile defense, with their models presented from a classical control point of view (see [3] [4]). A spiraling reentry vehicle analysis for Mars entry using an earlier version of the model in this thesis was presented by Dubois-Matra [5]. The motion model of the vehicle dynamics was based on a model presented by Bishop [6] for aircraft tracking. More recently, results have been reported on the estimation of the state of spiraling targets using the extended Kalman filters and unscented Kalman filters. In these cases, the model is presented without a physical basis for the parameters of the vehicle and atmosphere [7] [8].

Estimation techniques from a Bayesian point of view are well-known (see the detailed literature review in Chapter 3), and are useful for a wide variety of applications. Previous work shows the validity of the Bayesian framework for the estimation of the state of both linear and nonlinear systems. Aerospace applications have been the focus for much of the

modern development of these techniques, but Bayesian estimation is finding applications in a lot of scientific endeavours, for instance, economics [9], biological processes [10], stochastic optimal control [11] and robotics [12].

1.4 Contributions

This thesis has three main contributions. First, the mathematical model presented by [5] and Bishop and Antoulas [6] are advance by a parametrization that represents axisymmetric vehicles. Second, for the first time the spiraling frequency is estimated through the torsion using both the extended Kalman filter and the particle filter. Finally, the particle filter was successfully implemented for the first time for spiraling targets.

1.5 Thesis Organization

Chapter 2 presents the mathematical model of motion taking in account the atmospheric and gravity models. A parametrization is discussed upon which torsion is presented as a measure of the spiraling of the vehicle. Chapter 3 presents the theoretical basis of the Bayesian estimation framework, and a general form of the Bayes recursive filter. A general description of the estimation techniques under this framework is presented. Chapters 4 and 5 present the extended Kalman filter and the particle filter, respectively. Chapter 6 presents the results of applying the Bayesian techniques employing the mathematical model developed in Chapter 2 in realistic entry scenarios. Finally, Chapter 7 presents the conclusions and possible future work.

CHAPTER 2

Model Development and Analysis

In this chapter a mathematical model describing the motion of an axisymmetric reentry vehicle is presented. This model describes the translational dynamics and is physics based. In contrast, most existing models are artificially parameterized. The derivation of the mathematical model of motion is first presented, taking in account the atmosphere and gravity. A parametrization is proposed to provide structure to the model. The concept of torsion is discussed as a measure of the spiraling of the vehicle.

2.1 Model Development

2.1.1 Reference Frame Definitions

Consider the planet-centered, inertial frame shown in Fig 2.1 . The unit vector of the z-axis lies along the planet spin axis. The remaining two unit vectors lie in the planet equatorial plane and are oriented at a given epoch according to international agreements. The planet-centered, inertial reference frame is represented by ($\mathbf{u}_x$, $\mathbf{u}_y$, $\mathbf{u}_z$). Consider the unit vectors ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$) represent a rotating coordinate system attached to the reentry vehicle. Then, starting with Poisson's formula we have,

$$\dot{\mathbf{e}}_i = \boldsymbol{\omega} * \mathbf{e}_i \ , \quad \text{for } i = 1,\ 2, \text{or } 3 \ ,$$

where $\boldsymbol{\omega}$ is the angular velocity of the frame. It follows that

$$\mathbf{e}_i * \dot{\mathbf{e}}_i = \mathbf{e}_i * (\boldsymbol{\omega} * \mathbf{e}_i) = (\mathbf{e}_i \odot \mathbf{e}_i)\boldsymbol{\omega} - (\mathbf{e}_i \odot \boldsymbol{\omega})\mathbf{e}_i.$$

Since

$$\mathbf{e}_i \odot \mathbf{e}_i = 1 \ ,$$

we have the relationship

$$\boldsymbol{\omega} = (\mathbf{e}_i \odot \boldsymbol{\omega})\mathbf{e}_i + \mathbf{e}_i * \dot{\mathbf{e}}_i \qquad \text{for } i = 1,\ 2,\ 3 \ . \tag{2.1}$$

The relative velocity is $\dot{\mathbf{r}}_r = \dot{\mathbf{r}} - \boldsymbol{\Omega} * \mathbf{r}$, where the rotation of the planet $\boldsymbol{\Omega} = [0\ 0\ \Omega]^T$, and $\mathbf{r}$ is the position in the planet-centered, inertial reference frame. The wind frame reference vector $\mathbf{e}_1^w$ is defined as

$$\mathbf{e}_1^w = \frac{\dot{\mathbf{r}}_r}{\dot{r}_r} \ , \tag{2.2}$$

where $\dot{r}_r := \|\dot{\mathbf{r}}_r\|$ is the relative velocity magnitude of the vehicle, and $\dot{\mathbf{r}}$ is the inertial velocity in the planet-centered inertial reference frame, the superscript "w" denotes the wind reference frame, and the subscript "r" denotes relative to the rotating planet-centered, planet-fixed reference frame.

Since the vectors $\mathbf{e}_2^w$ and $\mathbf{e}_3^w$ only need to span the lift space, they can be the arranged to form a proper right-handed coordinate frame (see definitions in [13]), thus

$$\mathbf{e}_2^w = -\frac{\frac{\dot{\mathbf{r}}_r}{\dot{r}_r} * \mathbf{u}_z}{\left\| \frac{\dot{\mathbf{r}}_r}{\dot{r}_r} * \mathbf{u}_z \right\|} \qquad \text{and} \qquad \mathbf{e}_3^w = \mathbf{e}_1^w * \mathbf{e}_2^w \ , \tag{2.3}$$

where $\mathbf{u}_z = (\ 0 \quad 0 \quad 1\ )^T$. Then, $\mathbf{e}_2^w$ lies in a plane parallel to the inertial $x - y$ plane (that is, the equatorial plane of the planet). The transformation matrix from the wind frame to

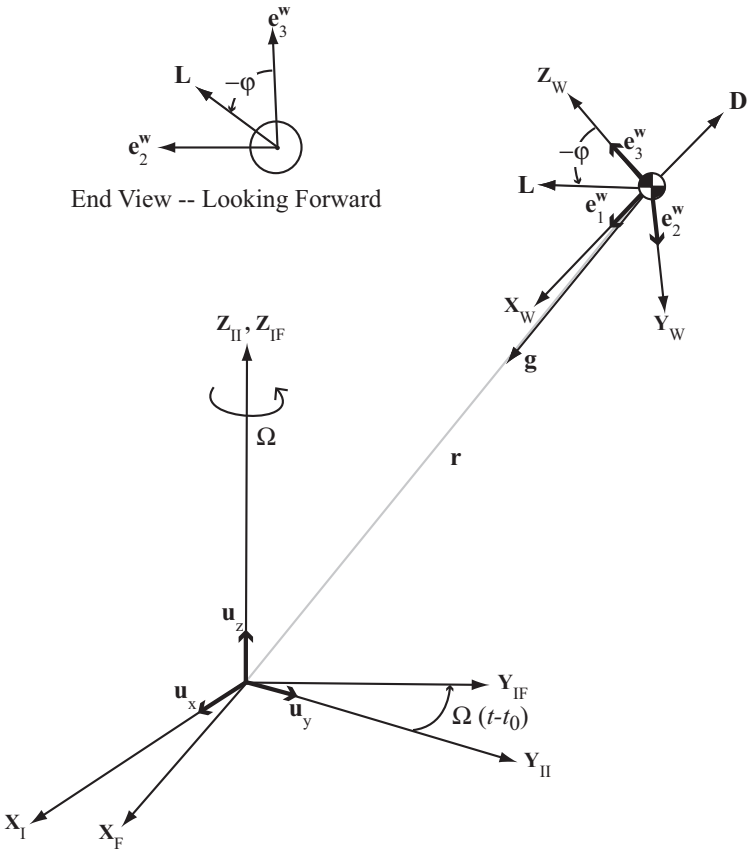

Figure 2.1: Entry vehicle model reference frames and lift/drag/bank angle definitions.

the inertial frame is given by

$$\mathbf{L}_{IW} = \left[\begin{array}{ccccc} \mathbf{e}_1^w & \vdots & \mathbf{e}_2^w & \vdots & \mathbf{e}_3^w \end{array} \right] .$$

Finally to find an expression for $\boldsymbol{\omega}$, recall that Eq. (2.1) is valid for any i =1, 2, or 3. For this development $i = 1$ is chosen. It can be said then that the term $\mathbf{e}_1 \odot \boldsymbol{\omega}$ in Eq. (2.1) is zero. This follows from the definition of the $\mathbf{e}_i^w$ vectors by constraining $\mathbf{e}_2^w$ to be in the equatorial plane. With $\mathbf{e}_1^w$ given in Eq. (2.2) and $\mathbf{e}_1 \odot \boldsymbol{\omega} = 0$, Eq. (2.1) reduces to

$$\boldsymbol{\omega}^w = \frac{\dot{\mathbf{r}}_r * \ddot{\mathbf{r}}_r}{\dot{r}_r^2} \ . \tag{2.4}$$

where $\ddot{\mathbf{r}}_r = \ddot{\mathbf{r}} - \boldsymbol{\Omega} * \dot{\mathbf{r}}$.

2.1.2 Vehicle Motion

Two key assumptions are made about the environment regarding the gravity and atmospheric density. First, we have a central Newtonian gravity field given by

$$\mathbf{g}(\mathbf{r}) = -\frac{\mu}{r^3}\mathbf{r} \ , \tag{2.5}$$

where μ is the constant planet gravity parameter, $\mathbf{r}$ is the position of the center of mass of the vehicle in the planet-centered, inertial reference frame, and $r = \|\mathbf{r}\|$. Second, we assume an exponential atmospheric density given by

$$\rho(r) = \rho_o e^{-(r-R_p)/H_o} \ . \tag{2.6}$$

where R_p is the planet radius (assuming a spherical planet), H_o is the base reference altitude associated with the density model assumed, and ρ_0 is the base reference density. These assumption are not critical to the mathematical model development and can be changed to include higher-order gravity and better high altitude atmospheric models, but are used here for the sake of simplicity.

The sum of the accelerations acting on the vehicle is given by

$$\ddot{\mathbf{r}} = \mathbf{a} + \mathbf{g} \ , \tag{2.7}$$

where $\mathbf{a}$ are the aerodynamic accelerations ($\mathbf{a} = \mathbf{L} + \mathbf{D}$ assuming the vehicle is axisymmetric, otherwise side accelerations $\mathbf{S}$ should be included) and $\mathbf{g}$ is the gravitational acceleration. We assume that the vehicle is not thrusting. Taking the time derivative of Eq. (2.7) yields

$$\dddot{\mathbf{r}} = \dot{\mathbf{a}} + \dot{\mathbf{g}} \ . \tag{2.8}$$

Using Eq. (2.5) to compute $\dot{\mathbf{g}}$ yields

$$\dot{\mathbf{g}} = -\frac{\mu}{r^3}\dot{\mathbf{r}} + 3\frac{\mu}{r^5}(\mathbf{r} \odot \dot{\mathbf{r}})\mathbf{r} \ . \tag{2.9}$$

Computing $\dot{\mathbf{a}}$ requires more effort. Referring to Fig. 2.1, it can be seen that

$$\mathbf{a} = -D\mathbf{e}_1^w + L\left[-\mathbf{e}_2^w \sin\varphi + \mathbf{e}_3^w \cos\varphi\right] \ , \tag{2.10}$$

where $D = \|\mathbf{D}\|$ and $L = \|\mathbf{L}\|$, and φ is the bank angle. Taking the time derivative of Eq. (2.10) yields

$$\dot{\mathbf{a}} = \left[\boldsymbol{\omega}^w + \dot{\varphi}\ \mathbf{e}_1^w\right] * (\ddot{\mathbf{r}} - \mathbf{g}) - \dot{D}\mathbf{e}_1^w + \dot{L}\left[-\mathbf{e}_2^w \sin\varphi + \mathbf{e}_3^w \cos\varphi\right] \ . \tag{2.11}$$

Substituting Eqs. (2.2), (2.3), (2.4), (2.9), and (2.11) into Eq. (2.8) yields

$$\begin{aligned} \dddot{\mathbf{r}} \ &= \ \overbrace{\left[\boldsymbol{\omega}^w * \left(\ddot{\mathbf{r}} + \frac{\mu}{r^3}\mathbf{r}\right)\right]}^{[a]} + \overbrace{\left[\dot{\varphi}\ \mathbf{e}_1^w * \left(\ddot{\mathbf{r}} + \frac{\mu}{r^3}\mathbf{r}\right)\right]}^{[b]} - \overbrace{\left[\dot{D}\mathbf{e}_1^w\right]}^{[c]} \\ &+ \ \underbrace{\dot{L}\left[-\mathbf{e}_2^w \sin\varphi + \mathbf{e}_3^w \cos\varphi\right]}_{[d]} - \underbrace{\frac{\mu}{r^3}\left[\dot{r}_r\mathbf{e}_1^w + \boldsymbol{\Omega} * \mathbf{r} - 3(\frac{\mathbf{r}}{r} \odot \dot{r}_r\mathbf{e}_1^w)\frac{\mathbf{r}}{r}\right]}_{[e]} \ . \end{aligned} \tag{2.12}$$

where

$$\boldsymbol{\omega}^w = \frac{\dot{\mathbf{r}}_r * \ddot{\mathbf{r}}_r}{\dot{r}_r^2} \ , \quad \mathbf{e}_1^w = \frac{\dot{\mathbf{r}}_r}{\dot{r}_r} \ , \quad \mathbf{e}_2^w = -\frac{\frac{\dot{\mathbf{r}}_r}{\dot{r}_r} * \mathbf{u}_z}{\left\|\frac{\dot{\mathbf{r}}_r}{\dot{r}_r} * \mathbf{u}_z\right\|} \ , \quad \text{and} \quad \mathbf{e}_3^w = \mathbf{e}_1^w * \mathbf{e}_2^w$$

From the expression in Eq. (2.12) several important characteristics of the vehicle motion can be recognized. The first term [a] represents curvature motion in-plane because $\boldsymbol{\omega}^w$ is by definition perpendicular to the maneuver plane. The value of $\boldsymbol{\omega}^w$ is related to the concept known as curvature which is a measure of the amount of turning in the maneuver plane [14]. The second term [b] represents out-of-plane motion due to rotation of the lift vector. When $\dot{\varphi} \neq 0$, the out-of-plane motion is nonzero. A similar situation with the fourth term [d], since it represents variations in the lift magnitude, which is a function of φ and $\dot{\varphi}$. The third term

[c] has no effect on the out-of-plane motion, but represents the variation of the drag force due to variations in velocity, atmospheric density and induced drag due to lift. The fifth term [e] represents the change in gravity and would only affect the out-of-plane motion when the maneuver plane is not vertical. When $\dot{\varphi} = 0$ (no rolling motion), $\mathbf{g}$ is a constant (i.e., a flat planet), $\dot{D} = \dot{L} = 0$, then the motion is confined to a plane—this is the so-called "coordinated turn model" (see [6]).

2.1.3 Model Parametrization

The lift and drag acceleration magnitudes are given by

$$L = \frac{\rho(r)\dot{r}_r^2 C_L S}{2m} \qquad \text{and} \qquad D = \frac{\rho(r)\dot{r}_r^2 C_D S}{2m} \ . \tag{2.13}$$

where C_L is the lift coefficient, C_D is the drag coefficient, $\rho(r)$ is the atmospheric density, S is the effective area of the vehicle, and m is the mass. The parameters C_L and C_D are usually not known at every moment during a vehicle reentry since they depend on Mach number, angle of attack and shape.

For the purpose of estimation, it is desired that the number of model parameters be small to reduce the computational complexity by minimizing the number of state variables. For a vehicle of general shape, the lift and drag coefficient are related in a near parabolic fashion. Figure 2.2 illustrates the parabolic relationship, known as the drag polar, that starts from C_{D_o}, the "zero-lift drag coefficient," that is, the drag coefficient when the vehicle is not generating lift, this for any given Mach number [15]. For the case of an axisymmetric vehicle (which is the case in this work), the drag polar is shown in Fig. 2.3. Let C_L^* denote the "critical lift coefficient," that is, the lift coefficient at maximum lift-to-drag ratio. The

drag coefficient is commonly modeled by the function

$$C_D = C_{D_o}\left[1 + \left(\frac{C_L}{C_L^*}\right)^2\right] , \tag{2.14}$$

where the term $(C_L/C_L^*)^2$ is the induced drag. In general, C_D can be represented by the function

$$C_D = C_{D_o} + KC_L^n ,$$

where K is a proportionality constant and n is to be determined for the particular vehicle.

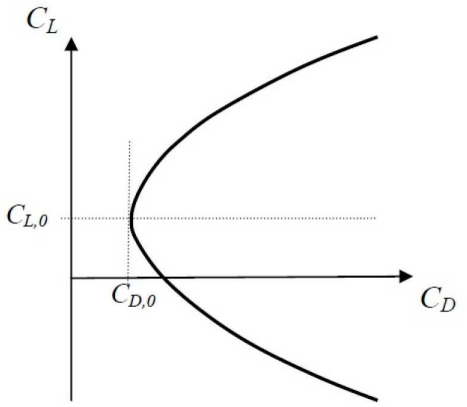

Figure 2.2: Drag polar for general shaped vehicle [15]

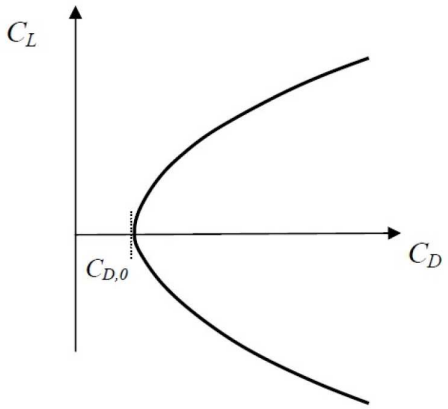

Figure 2.3: Drag polar for axisymmetric vehicle [15]

When using the more general form, it follows that

$$\frac{C_L}{C_D} = \frac{C_L}{C_{D_o} + KC_L^n} \ .$$

Taking the partial derivative with respect to C_L and setting the result to zero yields

$$K = \frac{C_{D_o}}{(n-1)(C_L^*)^n} \ .$$

Thus,

$$\left(\frac{C_L}{C_D}\right)_{\max} = \frac{(n-1)C_L^*}{nC_{D_o}} \ . \tag{2.15}$$

According to Regan and Anandakrishnan [13] it is usual to choose $n = 2.0$ for axisymmetric vehicles (for altitude between 0 and 100 km) leading to the relationship

$$C_{D_o} = \frac{C_L^*}{2\left(\frac{C_L}{C_D}\right)_{\max}} \ . \tag{2.16}$$

Using Eq. (2.15) with $n = 2$ and Eq. (2.16), we find that

$$C_L = 2C_{D_o}\left(\frac{C_L}{C_D}\right)_{\max}\left(\frac{C_L}{C_L^*}\right) \ , \tag{2.17}$$

and for convenience, the ratio between the lift coefficient and the critical lift coefficient is defined to be

$$\lambda := \frac{C_L}{C_L^*} \ . \tag{2.18}$$

With λ defined as in Eq. (2.18), the lift and drag magnitudes in Eq.(2.13), respectively, can be re-written as

$$\begin{aligned} L &= \left[2\beta_m\left(\frac{C_L}{C_D}\right)_{\max}\right]\lambda q(r,\dot{r}) \\ D &= \left[\beta_m(1+\lambda^2)\right] q(r,\dot{r}) \ , \end{aligned} \tag{2.19}$$

where $q(r, \dot{r})$ is the dynamic pressure given by

$$q(r, \dot{r}) = \frac{1}{2}\rho(r)\dot{r}_r^2 \ , \tag{2.20}$$

and the ballistic coefficient is defined to be

$$\beta_m := \frac{C_{D_o}S}{m} \ .$$

It is assumed that C_{D_o} remains essentially constant over the region of interest [13]. Taking the time derivative of C_D in Eq. (2.14) yields

$$\dot{C}_D = 2C_{D_o}\lambda\dot{\lambda} \ . \tag{2.21}$$

It is observed that when the vehicle is generating lift ($\lambda \neq 0$) and the lift magnitude (intentionally or unintentionally) varies ($\dot{\lambda} \neq 0$), the C_D will change due to the induced drag effects. Similarly, taking the derivative of C_L in Eq. (2.17) yields

$$\dot{C}_L = 2C_{D_o}\left(\frac{C_L}{C_D}\right)_{\max}\dot{\lambda} \ . \tag{2.22}$$

Other important derivatives include the density gradient (see Eq. (2.6))

$$\dot{\rho}(r) = -\rho(r)\frac{\dot{\mathbf{r}} \odot \mathbf{r}}{H_o r}, \tag{2.23}$$

and the relative velocity gradient

$$\frac{d(\dot{r}_r^2)}{dt} = 2\left[\ddot{\mathbf{r}} \odot \dot{\mathbf{r}} - (\boldsymbol{\Omega} * \mathbf{r}) \odot \ddot{\mathbf{r}}_r\right] \ . \tag{2.24}$$

Using Eq. (2.21)-(2.24) and taking the time-derivative of L and D in Eq.(2.19) yields

$$\begin{aligned} \dot{L} &= 2\beta_m\left(\frac{C_L}{C_D}\right)_{\max} q(r, \dot{r})\left[\dot{\lambda} + \lambda\left(\frac{2\ddot{\mathbf{r}} \odot \dot{\mathbf{r}}}{\dot{r}_r^2} - \frac{2(\boldsymbol{\Omega} * \mathbf{r}) \odot \ddot{\mathbf{r}}_r}{\dot{r}_r^2} - \frac{\dot{\mathbf{r}} \odot \mathbf{r}}{H_o r}\right)\right] \\ \dot{D} &= \beta_m q(r, \dot{r})\left[2\lambda\dot{\lambda} + (1+\lambda^2)\left(\frac{2\ddot{\mathbf{r}} \odot \dot{\mathbf{r}}}{\dot{r}^2} - \frac{2(\boldsymbol{\Omega} * \mathbf{r}) \odot \ddot{\mathbf{r}}_r}{\dot{r}_r^2} - \frac{\dot{\mathbf{r}} \odot \mathbf{r}}{H_o r}\right)\right] \ . \end{aligned} \tag{2.25}$$

When $\lambda \neq 0$, this implies that the vehicle is generating lift, and $\dot{\lambda} \neq 0$ implies that the vehicle is changing lift. The vehicle motion is described by specifying the control inputs $\varphi(t)$ and $\lambda(t)$ and the vehicle parameters β_m and $(C_L/C_D)_{\max}$. This parametrization give us the expressions that are needed in Eq. (2.12) to build the simulation of the vehicle motion and to create the state space used.

2.2 Spiraling Motion Analysis

Since the concern is the "spiraling" motion, it is desirable to quantify the spiraling motion in terms of the state of the vehicle. A measure of out-of-plane motion is the torsion. The torsion measures the rate at which the osculating plane turns as the vehicle moves along the trajectory [14]. When the torsion is zero, the motion is planar. The torsion is calculated via

$$\tau = \frac{\ddot{\mathbf{r}} * \dot{\mathbf{r}}}{\left\|\ddot{\mathbf{r}} * \dot{\mathbf{r}}\right\|^2} \odot \dddot{\mathbf{r}} \ ,$$

and has units of $1/length$. Now we assume that $\boldsymbol{\Omega} = \mathbf{0}_{1\times 3}$ for an easier analysis. Then, the contribution of terms [a] and [c] in Eq. (2.12) to the torsion is zero. Computing the contribution to the torsion from terms [b], [d] and [e] yields

$$\tau = \overbrace{-\dot{\varphi}\left[\frac{1}{\dot{r}} + \frac{\dot{r}}{||\ddot{\mathrm{r}} * \dot{\mathrm{r}}||^2}(\ddot{\mathrm{r}} - (\mathrm{e}_1^w \odot \ddot{\mathrm{r}})\mathrm{e}_1^w) \odot \frac{\mu}{r^3}\mathrm{r}\right]}^{\tau_b} - \overbrace{\dot{L}\left[\frac{\dot{r}}{||\ddot{\mathrm{r}} * \dot{\mathrm{r}}||^2}(\mathrm{e}_3^w \sin\varphi + \mathrm{e}_2^w \cos\varphi) \odot \mathrm{u_z}\right]}^{\tau_d} \underbrace{- \frac{3\mu}{r^3}\frac{\dot{r}^2}{||\ddot{\mathrm{r}} * \dot{\mathrm{r}}||^2}(\frac{\mathrm{r}}{r} \odot \mathrm{e}_1^w)\left[(\frac{\mathrm{r}}{r} * \ddot{\mathrm{r}}) \odot \mathrm{e}_1^w\right]}_{\tau_e} . \quad (2.26)$$

Here the contributors to the torsion are denoted by τ_b, τ_d, and τ_e, because of their relationship with the terms of the model in Eq. (2.12). The term τ_b is associated with the rotation rate $\dot{\varphi}$ of the lift vector, τ_d is related to the change in magnitude of the lift vector,

and τ_e is related to changes in the gravitational field.

With the objective of defining a "spiral frequency" that can be used to quantify the spiral motion, the torsion can be scaled by the velocity giving a measure of the rate (in standard units of frequency, such as rad/sec or 1/sec) of the motion of the osculating plane. Hence, the use of the term spiral frequency, which will be denoted by τ_s. Multiplying the torsion formula above by $\dot{r}$ and re-arranging terms leads to the following expression for the spiral frequency:

$$\tau_s = \dot{\varphi} - \frac{1}{\kappa^2}\frac{1}{\dot{r}^2}\left[\dot{\varphi}\left[\ddot{\mathbf{r}} - \left(\mathbf{e}_1^w \odot \ddot{\mathbf{r}}\right)\mathbf{e}_1^w\right] \odot \frac{\mu}{r^3}\mathbf{r} - \dot{L}\left(\sin\varphi\mathbf{e}_3^w + \cos\varphi\mathbf{e}_2^w\right) \odot \mathbf{u}_z \right. \\ \left. -\frac{3\mu}{r^3}\dot{r}\left(\frac{\mathbf{r}}{r} \odot \mathbf{e}_1^w\right)\left[\left(\frac{\mathbf{r}}{r} * \ddot{\mathbf{r}}\right) \odot \mathbf{e}_1^w\right]\right] , \tag{2.27}$$

where $\kappa = \|\omega^w\|$ is the curvature. As might be expected, the spiral frequency is directly proportional to $\dot{\varphi}$, which makes this value the most important factor that affects the change in torsion. However, the spiral frequency is also affected by other factors, including the curvature, lift variations, and gravity gradients (for maneuvers not in a vertical plane).

When $\Omega \neq 0$ there will be more terms in Eq. (2.26) that can even present influence from the drag magnitude in relationship with the rotation of the planet. Since the time of a reentry maneuver is in most applications much smaller than the rotation of a planet the effects added by these terms are negligible.

2.3 Model State Space Representation

The representation of the model presented in Eq. (2.12) is more useful when implementing estimation algorithms in a state space representation. A state space with $n = 11$ states was selected, where the states are the position, $\mathbf{r}$, velocity, $\dot{\mathbf{r}}$, acceleration, $\ddot{\mathbf{r}}$,

on the planet-centered, inertial reference frame, and the bank angle, φ, and the ratio of the lift coefficient to the critical lift coefficient, λ. Then the following vectors and variables can be defined $\mathbf{x}_1 = \mathbf{r} \in \mathbb{R}^{3\times1}$, $\mathbf{x}_2 = \dot{\mathbf{r}} \in \mathbb{R}^{3\times1}$, $\mathbf{x}_3 = \ddot{\mathbf{r}} \in \mathbb{R}^{3\times1}$, $x_4 = \lambda$, and $x_5 = \varphi$. Where the state vector is $\mathbf{x}(t) = \begin{bmatrix}\mathbf{x}_1^T & \mathbf{x}_2^T & \mathbf{x}_3^T & x_4 & x_5\end{bmatrix}^T$. Then establishing a system of differential equation we have

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \mathbf{x}_2 \\ \hline \mathbf{x}_3 \\ \hline \left[\boldsymbol{\omega}^w * \left(\mathbf{x}_3 + \frac{\mu}{\|\mathbf{x}_1\|^3}\mathbf{x_1}\right)\right] + \left[a\mathbf{e}_1^w * \left(\mathbf{x}_3 + \frac{\mu}{\|\mathbf{x}_1\|^3}\mathbf{x_1}\right)\right] \\ \quad - \left[\dot{D}\mathbf{e}_1^w\right] + \dot{L}\left[-\mathbf{e}_2^w \sin(x_5) + \mathbf{e}_3^w \cos(x_5)\right] \\ -\frac{\mu}{\|\mathbf{x}_1\|^3}\left[\|\mathbf{x}_{2_r}\|\mathbf{e}_1^w + \boldsymbol{\Omega} * \mathbf{x}_1 - 3(\frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} \odot \|\mathbf{x}_{2_r}\|\mathbf{e}_1^w)\frac{\mathbf{x}_1}{\|\mathbf{x}_1\|}\right] \\ \hline 0 \\ \hline a \end{bmatrix}, \tag{2.28}$$

where

$$\begin{aligned}
\dot{L} &= 2\beta_m \left(\frac{C_L}{C_D}\right)_{\max} q(\|\mathbf{x}_1\|, \|\mathbf{x}_2\|) x_4 \left(\frac{2\mathbf{x}_3 \odot \mathbf{x}_2}{\|\mathbf{x}_{2_r}\|^2} - \frac{2(\boldsymbol{\Omega} * \mathbf{x}_1) \odot \mathbf{x}_{3_r}}{\|\mathbf{x}_{2_r}\|^2} - \frac{\mathbf{x}_2 \odot \mathbf{x}_1}{H_o\|\mathbf{x}_1\|}\right), \\
\dot{D} &= \beta_m q(\|\mathbf{x}_1\|, \|\mathbf{x}_2\|)(1 + x_4^2) \left(\frac{2\mathbf{x}_3 \odot \mathbf{x}_2}{\|\mathbf{x}_{2_r}\|^2} - \frac{2(\boldsymbol{\Omega} * \mathbf{x}_1) \odot \mathbf{x}_{3_r}}{\|\mathbf{x}_{2_r}\|^2} - \frac{\mathbf{x}_2 \odot \mathbf{x}_1}{H_o\|\mathbf{x}_1\|}\right), \\
\mathbf{x}_{2_r} &= \mathbf{x}_2 - \boldsymbol{\Omega} * \mathbf{x}_1, \\
\mathbf{x}_{3_r} &= \mathbf{x}_3 - \boldsymbol{\Omega} * \mathbf{x}_2, \\
\boldsymbol{\omega}^w &= \frac{\mathbf{x}_{2_r} * \mathbf{x}_{3_r}}{\|\mathbf{x}_{2_r}\|^2}, \quad \mathbf{e}_1^w = \frac{\mathbf{x}_{2_r}}{\|\mathbf{x}_{2_r}\|}, \quad \mathbf{e}_2^w = -\frac{\mathbf{e}_1^w * \mathbf{u}_z}{\|\mathbf{e}_1^w * \mathbf{u}_z\|}, \quad \text{and} \quad \mathbf{e}_3^w = \mathbf{e}_1^w * \mathbf{e}_2^w.
\end{aligned}$$

The letter a represents a constant. It is important to note that the ballistic coefficient β_m (important since the vehicle could be unknown) can easily be transformed in one of the state variables and estimated.

CHAPTER 3

Bayesian Estimation

3.1 Introduction

We are often faced with decision making in the presence uncertainty. This uncertainty stems from situations where direct knowledge is not available, or to future predictions. Decisions are often based on inferences made from models of what is expected to be observed. Bayesian inference is the process of adjusting a probabilistic model to a set of data and summarizing the results through a probability density function (pdf) with the model parameters and the quantities that have not been observed. Bayesian estimation is the particularization of this concept to the filtering problem that consists of estimating the state of a system (physical, economic, etc) based on measurements that have a relationship with the states. Probability distributions are used for modeling both the uncertainties in the system model and parameters, and for modeling the characteristics of the random elements of the system.

Bayesian inference is a theoretical, yet practical framework for reasoning, decision making and estimation under uncertainty. The historical roots of the theory lie in the late 18th and early 19th century with Thomas Bayes and Pierre-Simon de Laplace [16]. Bayesian inference was not a popular approach for decision making until the last half of the 20th century. Bayesian inference did not develop as a single, homogeneous scientific activity. It

has, however, been employed in many different domains. The Bayesian approach to filtering is not new (see, e.g., [16]; [17] ; [18]). The theory appeared in the seminal article of Kalman [19]. The Kalman filter can be derived from the mean least-squares point of view (optimization) or from a Bayesian perspective. Non-linear filtering theory, such as the extended Kalman filter (EKF), are generally Bayesian from the beginning (see, e.g., [17]). As computations became faster and more accessible, state estimators with a higher computational cost were developed. From these estimators we consider three categories. First, we consider the class of different variants of the EKF that provide estimates of the state variables, and a measure of the mean least-square state estimation error. In this category we can include the estimators that approximate the pdf of the variable with a mixture of probability density functions. This was first proposed by Sorenson and Alspach [20], using a mixture of Gaussians. We can also consider grid based filters that evaluate the pdf using a series of nodes chosen to cover the entire state space. This set of nodes, each with an associated weight, are used as a discrete approximation of the posterior pdf or as base for continuous approximations for this pdf, for example using splines [21]. The last category of filters are those that use Monte Carlo methods. Their origins can be traced to Handschin [22] and Mayne and Handschin [23]. Gordon et al [24] employ sequential Monte Carlo methods as set of points that approximate the posterior pdf.

In this chapter a derivation of the general recursive Bayesian filter is presented. First, important concepts related with Bayesian filtering are discussed. Next, the concepts are used to show the development of Bayesian estimation and the general recursive algorithm. Finally, a general description of the different Bayesian filters is presented.

3.2 Bayesian Inference

A scientific hypothesis is typically represented as a pdf of the observed data. This pdf depends on certain unknown quantities or parameters, denoted by θ. In the Bayesian paradigm, the knowledge of the model parameters are expressed through a pdf, known as the prior density function, $p(\theta)$. When new data y is obtained, the information contained in the prior pdf and its relation with the model parameters is known as the "likelihood" function, and is represented by $p(y|\theta)$. The information contained in the prior pdf and the likelihood function can be combined to obtain a new pdf, known as the posterior pdf and denoted by $p(\theta|y)$. The posterior pdf is the objective of the Bayesian inference process. Bayes theorem is an elemental identity in probability theory (more information on this and basic probability theory can be found in [25]). According to Bayes, the posterior probability is proportional to the product of the priori by the likelihood,

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{\int p(\theta)p(y|\theta)\mathrm{d}\theta}.$$

In theory, one can always obtain the posterior distribution, but with the complex systems and models the necessary analytical calculation are typically intractable. In recent years, the research community realized that obtaining samples of the posterior could be an applicable and adequate option.

There are several reasons to use Bayesian methods, and their applications are present in several different fields. Many investigations into the use of Bayesian methods have been reported. It is evident that if one wants to make a consistent decision in the presence of uncertainty, an excellent approach is to use Bayesian methods.

3.3 Continuous-Discrete Probabilistic Dynamical Systems

A probabilistic dynamical systems is a sequence of continuous probability density functions $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k})$, where $\mathbf{x}(t_k)$ is the state vector, $\mathbf{y}_k$ is the observations vector, the index $t = t_k$ represents an instant of time when a observation is obtained and the subscript $1:k$ represents the set of observation at all instants up to and including t_k. The state variable $\mathbf{x}(t)$ evolves over time. In most of the applications, the difference between $p(\bullet|\mathbf{y}_{1:k})$ and $p(\bullet|\mathbf{y}_{0:k-1})$ is due to the incorporation of a new observation. The following processes are of special interest:

Prediction: $p(\mathbf{x}(t+dt)|\mathbf{y}_{1:k}), dt \neq 0$, where $p(\bullet|\mathbf{y}_{1:k})$ can be computed for all time $t > t_k$. The best prediction of $\mathbf{x}(t)$ before new information arrives is through $p(\bullet|\mathbf{y}_{1:k-1})$.

Smoothing: $p(\mathbf{x}(t)|\mathbf{y}_{1:T}), 0 < t < t_T$. In this case, the distribution can be calculated for all times $t \in [0, t_T]$ if the observations up to the instance $\mathbf{y}_T$ have been observed.

Estimation: $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k})$ (Sequential estimation). Here we estimate the variable $\mathbf{x}(t_k)$ at the time instance t_k when the observation $\mathbf{y}_k$ has been obtained.

For this work, we consider dynamical systems that are represented in a state space form. A state space model is defined by the state equation,

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) + \sigma(\mathbf{x}(t), t)\mathbf{w}(t), \tag{3.1}$$

and the measurement equation,

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}(t_k), t_k, \mathbf{v}(t_k)) \tag{3.2}$$

where $\mathbf{y}_k$ is the observations vector of t_k, $\mathbf{x}(t)$ is the state vector, $\mathbf{h}$ is the measurement function (vector of functions), $\mathbf{f}$ is the state function (also known as the drift function), $\mathbf{u}(t)$

is the vector of inputs or control actions, $\mathbf{w}(t)$ is a stochastic noise process, $\mathbf{v}(t_k)$ is a random noise sequence, t is the time dependance and t_k represents the instant an observation is obtained. The usual assumptions are that the analytical representation of the functions and distributions of both noises are known. The objective of Bayesian estimation in this case is to estimate $\mathbf{x}(t_k)$ in a recursive form based on the observations $\mathbf{y}_k$, obtaining the posterior distribution $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k})$.

State variables and measurements are directly related to the different probability density functions that represent the system when it is treated as a set of stochastic processes, and that are ultimately used for Bayesian estimation. In general, it can be said that

$$x \sim p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}))$$

$$y \sim p(\mathbf{y}_k|\mathbf{x}(t_k)),$$

where $p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}))$ is known as the transition density. There are two final definitions and assumptions that are key to Bayesian estimation and inference. First, the Markov assumption which states that the values in any state $\mathbf{x}(t)$ are only influenced by the values of the state $\mathbf{x}(t - dt)$ that directly preceded it. This implies that the past is independent of the future. In a continuous-discrete setting, we have

$$p(\mathbf{x}(t_{0:k})) = \prod_{i=1}^{k} p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}))p(\mathbf{x}(t_0)). \tag{3.3}$$

We also have the conditional independence of observations that states that the observation, $\mathbf{y}_k$, given the state, $\mathbf{x}(t_k)$, is conditionally independent from the observation and state

history, or

$$\begin{aligned} p(\mathbf{y}_{1:k}) &= \prod_{i=1}^{k} p(\mathbf{y}_i) \\ p(\mathbf{y}_{1:k}|\mathbf{x}(t_{0:k})) &= \prod_{i=1}^{k} p(\mathbf{y}_i|\mathbf{x}(t_i)). \end{aligned} \tag{3.4}$$

3.4 Recursive Estimation

The Bayesian filters considered here, namely the extended Kalman filter and the particle filter are based on a general structure. Each filter differs under different assumptions. The main objective in each case is to estimate the state of a system using observations, where the state evolves in the presence of noise and observations are made sequentially also in the presence of noise. The notation is as follows: $\mathbf{x}(t)$ is the state being estimated, and $\mathbf{y}_k$ indicates the observed data. The problem consists of estimating the state $\mathbf{x}(t_{0:k}), k = 1, 2, ...$ based on the sequence of observations $\mathbf{y}_{1:k}, k = 2, 3, ...$ In this derivation, the Markov assumption and conditional independence of observations assumption apply.

The set of posterior distributions can be represented using the Bayes theorem as

$$p(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_{1:k}|\mathbf{x}(t_{0:k}))p(\mathbf{x}(t_{0:k}))}{p(\mathbf{y}_{1:k})}. \tag{3.5}$$

In a practical setting all the information needed to compute $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k})$ is not known or cannot be obtained in real-time, so using the assumptions from Eqns. (3.3) and (3.4), we begin by rewriting Eq. (3.5) as

$$p(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k}) = \prod_{i=1}^{k} \frac{p(\mathbf{y}_i|\mathbf{x}(t_i))p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}))p(\mathbf{x}(t_0))}{p(\mathbf{y}_i)}. \tag{3.6}$$

Eq. (3.6) can be expanded sequentially to obtain an expression for $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k})$ by induction. So, expanding for $i = 1$ we have

$$p(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k}) = \prod_{i=2}^{k} \frac{p(\mathbf{y}_i|\mathbf{x}(t_i))p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}))}{p(\mathbf{y}_i)} \frac{p(\mathbf{y}_1|\mathbf{x}(t_1))p(\mathbf{x}(t_1)|\mathbf{x}(t_0))p(\mathbf{x}(t_0))}{p(\mathbf{y}_1)} \tag{3.7}$$

Integrating both sides of Eq. (3.7) with respect to $\mathbf{x}(t_0)$ gives

$$p(\mathbf{x}(t_{1:k})|\mathbf{y}_{1:k}) = \prod_{i=2}^{k} \frac{p(\mathbf{y}_i|\mathbf{x}(t_i))p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}))}{p(\mathbf{y}_i)} \underbrace{\frac{p(\mathbf{y}_1|\mathbf{x}(t_1))p(\mathbf{x}(t_1))}{p(\mathbf{y}_1)}}_{=p(\mathbf{x}(t_1)|\mathbf{y}_1)\ by\ Bayes}, \tag{3.8}$$

since

$$\begin{aligned} \int p(\mathbf{x}_{0:k})d\mathbf{x}(t_0) &= \prod_{i=2}^{k} p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1})) \underbrace{\int p(\mathbf{x}(t_1)|\mathbf{x}(t_0))p(\mathbf{x}(t_0))d\mathbf{x}(t_0)}_{p(\mathbf{x}(t_1))} \\ &= p(\mathbf{x}(t_{1:k})). \end{aligned}$$

Continuing for $i = 2$ we have

$$p(\mathbf{x}(t_{1:k})|\mathbf{y}_{1:k}) = \prod_{i=3}^{k} \frac{p(\mathbf{y}_i|\mathbf{x}(t_i))p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}))}{p(\mathbf{y}_i)} \underbrace{\frac{p(\mathbf{y}_2|\mathbf{x}(t_2))p(\mathbf{x}(t_2)|\mathbf{x}(t_1))p(\mathbf{x}(t_1)|\mathbf{y}_1)}{p(\mathbf{y}_2)}}_{p(\mathbf{x}(t_{1:2})|\mathbf{y}_{1:2})}. \tag{3.9}$$

Integrating with respect to $\mathbf{x}(t_1)$ in Eq. (3.9) yields

$$p(\mathbf{x}(t_{2:k})|\mathbf{y}_{1:k}) = \prod_{i=3}^{k} \frac{p(\mathbf{y}_i|\mathbf{x}(t_i))p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}))}{p(\mathbf{y}_i)} \underbrace{\frac{p(\mathbf{y}_2|\mathbf{x}(t_2))p(\mathbf{x}(t_2)|\mathbf{y}_1)}{p(\mathbf{y}_2)}}_{p(\mathbf{x}(t_2)|\mathbf{y}_{1:2})}, \tag{3.10}$$

since

$$p(\mathbf{x}(t_2)|\mathbf{y}_1) = \int p(\mathbf{x}(t_2)|\mathbf{x}(t_1))p(\mathbf{x}(t_1)|\mathbf{y}_1)d\mathbf{x}(t_1). \tag{3.11}$$

After expanding for the k^{th} instant and integrating sequentially for $\mathbf{x}(t_{k-1})$, we obtain

$$p(\mathbf{x}(t_k)|\mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k|\mathbf{x}(t_k))p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1})}{p(\mathbf{y}_{1:k})}, \tag{3.12}$$

where

$$p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1}) = \int p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}))p(\mathbf{x}(t_{k-1})|\mathbf{y}_{1:k-1})d\mathbf{x}(t_{k-1}) \tag{3.13}$$

Eq. (3.12) is the general form of the recursive Bayesian filter. The likelihood function $p(\mathbf{y}_k|\mathbf{x}(t_k))$, that represents the pdf of the observations depends on the noise of the sensor. The posterior pdf before a new observation is made is given by $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1})$. Eq. (3.13) is known as the Chapman-Kolmogorov equation.

3.4.1 Algorithm

After having all the elements that form the Bayesian recursive filter, we have that the algorithm is a recursive process that starts with $p(\mathbf{x}(t_0))$, the pdf associated with $\mathbf{x}(t)$ prior to any observations. The recursive algorithm is divided in two main steps, prediction and update, that are applied when each observation $\mathbf{y}_k$ is obtained.

The prediction step is where the pdf prior to an observation, $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1})$ is calculated. The continuous nature of the system is significant since a stochastic differential equation has to be solved. Theoretically, there are several ways to proceed. There is not an

unique approach to solve the stochastic differential equations, but also due to the complexity of the models of the system, most of the methods are in general intractable and not suitable for practical applications.

First Method: Propagate the transition density function $p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}))$ from time $t_{k-1} < t < t_k$ by integrating the stochastic differential equation that represent the state $\mathbf{x}(t)$ from time $t_{k-1} < t < t_k$. Using this result, calculate $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1})$ using Eq. (3.13). This method is typically computationally intractable, but can be approximated under some assumptions [26].

Second Method: Solve the boundary problem of finding $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1})$ starting from the distribution $p(\mathbf{x}(t_{k-1})|\mathbf{y}_{1:k-1})$ and solving the partial differential equation from time $t_{k-1} < t < t_k$. It is necessary to use numerical approximations in most cases.

For the update step, we compute the posterior pdf using Bayes theorem to incorporate the observation pdf where,

$$p(\mathbf{x}(t_k)|\mathbf{y_k}) \propto p(\mathbf{y}_k|\mathbf{x}(t_k))p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1})$$

As mentioned before, this is a general form of the Bayesian estimation. This exact structure will not be readily apparent in most filters, even though, in general, the prediction and update form is followed.

3.4.2 Bayesian Point Estimates and Optimal Filtering

In applied estimation situations, the use of complete pdf's is not necessary (and generally intractable), since depending on the assumptions made for a given filter, only a few parameters of a pdf need to be estimated. The most common point estimators used in

Bayesian estimation are the expected value and possibly some higher moments of the pdf.

A point estimate of a variable $\mathbf{x}$ is usually represented by $\hat{\mathbf{x}}$ typically represents the expected values or mean. It is also important to consider that the state estimates have to be considered stochastic processes as well. For example this common point estimator could be represented as the mean,

$$E(\mathbf{x}(t_{k|k})) = \hat{\mathbf{x}}(t_{k|k}) = \int \mathbf{x}(t_k) p(\mathbf{x}(t_k)|\mathbf{y}(t_k)) \mathrm{d}\mathbf{x}(t).$$

or mode

$$\hat{\mathbf{x}}(t_{k|k}) = max_{\mathbf{x}(t_k)} p(\mathbf{x}(t_k)|\mathbf{y}(t_k)).$$

There is also the important element added to the point estimators (explicit or implicitly) that is the loss function that defines a penalty for erroneous estimates. This is where the relationship between Bayesian estimation and optimization theory is found. An example is the Kalman filter that was first derived from a stochastic optimal control point of view. A common loss function $L(\mathbf{x}, \hat{\mathbf{x}})$ used in the continuous-discrete scenario could be,

$$L(\mathbf{x}, \hat{\mathbf{x}}) = \int_0^{t_T} (\mathbf{x}(t) - \hat{\mathbf{x}}(t))^T (\mathbf{x}(t) - \hat{\mathbf{x}}(t)) dt$$

The use of a loss function is not usually explicit for Bayesian point estimators, but they introduce the concept of optimality, with the objective of obtaining unbiased estimates.

3.5 Algorithms for Optimal Filtering

Starting from the general Bayesian recursive filter, and using different assumptions and systems, a wide variety of filters can be obtained. There are two different established

families of filters: point estimators (with Gaussian noise assumptions) and Monte Carlo methods for density estimation. All these techniques have versions in the different time representations (continuous-discrete in our case), but this does not affect their general characteristics except in how some steps are performed and some values are calculated. It is important to note that many filters strategies are not optimal, since usually approximations are employed.

3.5.1 Kalman Filter

The Kalman filter is an optimal point estimator where the two first moments of the posterior distribution are calculated in a recursive fashion [19]. It requires the system and observation models to be linear but not necessarily time-invariant. It is generally assumed that the process and measurement noises are Gaussian, which translates into the assumption that the prior and likelihood functions are Gaussian.

Due to the fact that a linear system is assumed and thanks to the special properties of Gaussian distribution functions, a posterior of the same kind is obtained with a linear transformation. This assures the preservation of optimality. More details about the mathematical process to obtain this filter within Bayesian framework will be shown in the next chapter.

3.5.2 Extended Kalman Filter

The extended Kalman filter is one the most widely used point estimator. The EKF does not make assumptions of linearity in the system and observation models and has the same structure as the Kalman filter. However the system and observation models must be

sufficiently differentiable, since the EKF uses Taylor series approximations requiring Jacobian matrices.

In the EKF derivation, higher order terms of the Taylor series are neglected making the EKF a linear approximation. The posterior represents an approximation to the true pdf. This makes the EKF a suboptimal filter. This filter is the defacto standard and serves as comparison for other non-linear estimation techniques.

3.5.3 Sigma Point Kalman Filter

The Sigma Point Kalman Filter (SPKF) handles nonlinear systems, but not through a analytic approximation as with the EKF, but instead using a minimal set of deterministically chosen weighted sample point (sigma points) that capture the true nature of the first and second moment of a Gaussian distribution. These sigma points are obtained through a deterministic transformation. Like the Kalman filters, all pdf's are assumed to be Gaussian, and like the extended Kalman filter the posterior is an approximation and is suboptimal.

There are many varieties of SPKF algorithms. The most established SPKF are the unscented Kalman filter [27], central-difference Kalman filter [28] and the Gauss-Hermite Kalman Filter [29], but there are more in the literature (see e.g., [30];[31]). An excellent paper that shows the derivation of the unscented Kalman filter for continuous-discrete systems from a Bayesian point of view can be found in [32].

3.5.4 Particle Filter

The particle filter is not related to the family of Kalman filters. The particle filter employs a different approach based on simulation (Monte Carlo methods) and sampling theory (importance sampling). It is not a point estimator but rather a pdf estimator, where an approximation of the "true" posterior pdf is obtained based on weighted samples. Then given this obtained posterior, the different moments can be calculated.

The particle filer has a similar structure to the general recursive Bayesian estimator described before, but non-parametric pdf's are obtained. There exist several variants depending on the assumptions, optimality needs and sampling or resampling techniques used. The most widely used particle filter is known as the sequential importance sampling particle filter. More details on this filter will be discussed in Chapter 5. The particle filter is chosen for this work because it does not require approximations of the nonlinear model. This is important when quantifying the influence of approximations in the accuracy of the state estimation for this problem.

CHAPTER 4

Extended Kalman Filter

A derivation of the extended Kalman filter (EKF) is presented along with the implementation details for this project. First, important concepts and definitions pertinent to the derivation are presented. Then using this information, a derivation of the continuous-discrete EKF from the Bayesian point of view is developed. Finally, all the elements related with the implementation are discussed.

4.1 Preliminary Concepts

4.1.1 Solution of Linear Stochastic Differential Equations with Gaussian Noise

Consider the system (in Chapter 3) given by Eq. (3.1) in the case where $f(\mathbf{x}(t), \mathbf{u}(t), t) = \mathbf{A}(t)\mathbf{x}(t)$ is a linear function, and $\mathrm{d}\mathbf{w}(t)$ is white noise process. Then, it follows that

$$\frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t} = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{w}(t) \tag{4.1}$$

is obtained, where $\mathbf{x}(t)$ is an n-vector, $\mathbf{w}(t)$ has a covariance $E[\mathbf{w}(t)\mathbf{w}^T(\tau)] = \mathbf{Q}_1\delta(t-\tau)$. The elements of $\mathbf{A}(t)$ and $\mathbf{Q}_1(t)$ are continuous functions of time. It is also assumed that $\mathbf{x}(t_0)$ is also a normally distributed random variable with an expected value $\hat{\mathbf{x}}_0 = E[\mathbf{x}(t_0)]$

and covariance $E[\hat{\mathbf{x}}_0\hat{\mathbf{x}}_0^T] = \mathbf{Q}_0$. A solution of Eq. (4.1) can be written as

$$\mathbf{x}(t) = \mathbf{\Phi}(t;t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t;\tau)\mathbf{w}(\tau)\mathrm{d}\tau \tag{4.2}$$

where $\mathbf{\Phi}(t;t_0)$ is the state transition matrix and satisfies the differential equation

$$\frac{d\mathbf{\Phi}(t;t_0)}{dt} = \mathbf{A}(t)\mathbf{\Phi}(t;t_0) \tag{4.3}$$

with the initial condition $\mathbf{\Phi}(t_0;t_0) = \mathbf{I}$. Since $\mathbf{x}(t)$ is a linear function of a Gaussian process, it is also Gaussian and can be characterized completely by the expected value (first moment of the pdf) and the covariance. Computing $E[\mathbf{x}(t)]$ yields

$$E[\mathbf{x}(t)] = \mathbf{\Phi}(t;t_0)E[\mathbf{x}(t_0)] + E\left[\int_{t_0}^{t} \mathbf{\Phi}(t;\tau)\mathbf{w}\mathrm{d}(\tau)\right]. \tag{4.4}$$

Since $\mathbf{w}(t)$ is zero-mean white noise, we have $E[\mathbf{w}(t)] = 0$, so the second term on the right side of Eq. (4.4) will vanish. Then, it follows that

$$\hat{\mathbf{x}}(t) = E[\mathbf{x}(t)] = \mathbf{\Phi}(t;t_0)E[\mathbf{x}(t_0)] = \mathbf{\Phi}(t;t_0)\hat{\mathbf{x}}_0. \tag{4.5}$$

Taking the derivative with respect to time in Eq. (4.5) yields

$$\frac{d\hat{\mathbf{x}}(t)}{dt} = \frac{d}{dt}\mathbf{\Phi}(t;t_0)\hat{\mathbf{x}}_0 = \mathbf{A}(t)\mathbf{\Phi}(t;t_0)\hat{\mathbf{x}}(t_0) = \mathbf{A}(t)\hat{\mathbf{x}}(t). \tag{4.6}$$

The initial value for Eq. (4.6) can be seen in Eq. (4.5) and is given by

$$\hat{\mathbf{x}}(t_0) = \hat{\mathbf{x}}_0. \tag{4.7}$$

To compute the state estimation error covariance we subtract $\hat{\mathbf{x}}(t)$ from $\mathbf{x}(t)$) to obtain

$$\mathbf{e}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t) = \mathbf{\Phi}(t;t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t;\tau)\mathbf{w}(\tau)\mathrm{d}\tau - \mathbf{\Phi}(t;t_0)\hat{\mathbf{x}}_0. \tag{4.8}$$

Then, compute $\mathbf{P}(t) = E[\mathbf{e}(t)\mathbf{e}^T(t)]$ yields

$$\mathbf{P}(t) = E[\mathbf{x}(t)\mathbf{x}^T(t)] = \mathbf{\Phi}(t;t_0)\mathbf{Q}_0\mathbf{\Phi}(t;t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t;\tau)\mathbf{Q}_1(\tau)\mathbf{\Phi}^T(t;\tau)\mathrm{d}\tau. \tag{4.9}$$

To compute the evolution in time of $\mathbf{P}(t)$, take derivative of $\mathbf{P}(t)$ in Eq. (4.9), yielding

$$\dot{\mathbf{P}}(t) = \mathbf{A}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^T(t) + \mathbf{Q}_1(t) \tag{4.10}$$

with

$$\mathbf{P}(t_0) = \mathbf{P}_0. \tag{4.11}$$

It is also important to note that the solution of Eq. (4.10) represents the prediction (also known as propagation) step of a continuous-discrete Kalman filter.

To summarize, we have

$$\dot{d\hat{\mathbf{x}}}(t) = \mathbf{A}(t)\hat{\mathbf{x}}(t)$$

$$\hat{\mathbf{x}}(t_0) = \hat{\mathbf{x}}_0$$

$$\dot{d\mathbf{P}}(t) = \mathbf{A}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^T(t) + \mathbf{Q}_1(t)$$

$$\mathbf{P}(t_0) = \mathbf{P}_0$$

4.1.2 Special Properties of Gaussian Distributions

The set of special properties and general characteristics of Gaussian distributions are directly related with the update step in a Kalman type Bayesian filter and can be obtained using regular calculus and probability definitions [25] [33] [34]. A random variable $\mathbf{x} \in \mathbb{R}^n$ has a Gaussian distribution with mean $\mu \in \mathbb{R}^n$ and covariance $\mathbf{P} \in \mathbb{R}^{n \times n}$ with the pdf of the form

$$N(\mathbf{x}|\mu, \mathbf{P}) = \frac{1}{(2\pi)^{n/2}|\mathbf{P}|^{1/2}} exp\bigg(-\frac{1}{2}(\mathbf{x}-\mu)^T\mathbf{P}^{-1}(\mathbf{x}-\mu)\bigg), \tag{4.12}$$

where $|\mathbf{P}|$ is the determinant of the matrix $\mathbf{P}$.

Joint density of Gaussian variables. If random variables $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$ have the Gaussian densities

$$\begin{aligned} \mathbf{x} &\sim N(\mathbf{x}|\mu, \mathbf{P}) \\ \mathbf{y}|\mathbf{x} &\sim N(\mathbf{y}|\mathbf{H}\mathbf{x}+\mathbf{u}, \mathbf{R}), \end{aligned} \tag{4.13}$$

where $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{H} \in \mathbb{R}^{m \times n}$, both independent of $\mathbf{x}$. Then the joint density of $\mathbf{x}, \mathbf{y}$ and the marginal distribution of $\mathbf{y}$ are given as

$$\begin{aligned} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} &\sim N\left(\begin{bmatrix} \mu \\ \mathbf{H}\mu + \mathbf{u} \end{bmatrix}, \begin{bmatrix} \mathbf{P} & \mathbf{P}\mathbf{H}^T \\ \mathbf{H}\mathbf{P} & \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R} \end{bmatrix} \right) \\ \mathbf{y} &\sim N(\mathbf{y}|\mathbf{H}\mu + \mathbf{u}, \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}) \end{aligned} \tag{4.14}$$

Conditional density of Gaussian variables. If the random variables $\mathbf{x}$ and $\mathbf{y}$ have the

joint Gaussian probability density

$$\mathbf{x}, \mathbf{y} \sim N\left(\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}, \begin{bmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^T & \mathbf{B} \end{bmatrix} \right) \tag{4.15}$$

then the marginal and conditional densities of $\mathbf{x}$ and $\mathbf{y}$ are given as follows:

$$\begin{aligned} \mathbf{x} &\sim N(\mathbf{a}, \mathbf{A}) \\ \mathbf{y} &\sim N(\mathbf{b}, \mathbf{B}) \\ \mathbf{x}|\mathbf{y} &\sim N(\mathbf{a} + \mathbf{C}\mathbf{B}^{-1}(\mathbf{y} - \mathbf{b}), \mathbf{A} - \mathbf{C}\mathbf{B}^{-1}\mathbf{C}^T) \\ \mathbf{y}|\mathbf{x} &\sim N(\mathbf{b} + \mathbf{C}^T\mathbf{A}^{-1}(\mathbf{x} - \mathbf{a}), \mathbf{B} - \mathbf{C}^T\mathbf{A}^{-1}\mathbf{C}). \end{aligned} \tag{4.16}$$

4.2 Continuous-Discrete Extended Kalman Filter

Like other Bayesian filters, the continuous-discrete EKF has a prediction/update structure. First, we consider $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1})$ starting from the system shown in Eq. (3.1) and using the procedure from Section 4.1.1. Then the update step is derived using the properties from Section 4.1.2, obtaining a Gaussian distribution equivalent to $p(\mathbf{x}(t_k)|\mathbf{y}_{1:k})$.

We assume a model of the form

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(\mathbf{x}(t), t) + \boldsymbol{\xi}(t) \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}(t_k), t_k) + \mathbf{v}_k \end{aligned} \tag{4.17}$$

where $\boldsymbol{\xi}$(t) is a stochastic noise process with $E[\boldsymbol{\xi}(t)\boldsymbol{\xi}^T(\tau)] = \mathbf{Q}(t)\delta(t - \tau)$ and $\mathbf{v}_k$ is a random noise sequence with $E[\mathbf{v}_k\mathbf{v}_j^T] = \mathbf{R}_k\delta_{kj}$

4.2.1 Prediction

Assume that the $\mathbf{f}(\mathbf{x}(t), t)$ in Eq. (4.17) is a nonlinear differentiable vector-valued function. This function can be expanded around a point using a Taylor series expansion. In the case of the EKF, the reference is the expected value of the state variable $\hat{\mathbf{x}}(t)$. Here, we have

$$\mathbf{f}(\mathbf{x}(t), t) = \mathbf{f}(\hat{\mathbf{x}}(t), t) + \mathbf{F}(\hat{\mathbf{x}}(t), t)(\mathbf{x}(t) - \hat{\mathbf{x}}(t)) + ... \tag{4.18}$$

Neglecting the higher order terms of the expansion yields

$$\dot{\mathbf{x}}(t) = \left[\mathbf{f}(\hat{\mathbf{x}}(t), t) + \mathbf{F}(\hat{\mathbf{x}}(t), t)(\mathbf{x}(t) - \hat{\mathbf{x}}(t))\right] + \boldsymbol{\xi}(t) \tag{4.19}$$

where $\mathbf{F}(\hat{\mathbf{x}}(t), t)$ is the Jacobian matrix,

$$\mathbf{F}(\hat{\mathbf{x}}(t), t) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \Bigg|_{\mathbf{x}(t)=\hat{\mathbf{x}}(t)}.$$

Another important expression can be observed in Eq. (4.19) is that $\mathbf{x}(t) - \hat{\mathbf{x}}(t)$ represents the estimation error $\mathbf{e}(t)$. In general the expected value of the estimation error is desired to be zero, $E[\mathbf{e}(t)] = 0$, so that the estimator is unbiased. Taking the expectation of Eq. (4.19) yields

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t), t) \quad t_{k-1} \leq t \leq t_k, \tag{4.20}$$

where $\hat{\mathbf{x}}(t_{k-1}) = \mathbf{x}_{k-1}^{+}$ is the state estimate at t_{k-1} after the measurement update in the previous iteration. The state estimation error covariance is found by solving

$$\dot{\mathbf{P}} = \mathbf{F}(\hat{\mathbf{x}}(t), t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}(\hat{\mathbf{x}}(t), t)^T + \mathbf{Q}(t) \qquad t_{k-1} \leq t \leq t_k, \tag{4.21}$$

with the initial condition

$$\mathbf{P}_0 = \mathbf{P}^+_{k-1}.$$

where $\mathbf{P}^+_{k-1}$ is the state estimate error covariance at t_{k-1} after the measurement update in the previous iteration. The main objective in the prediction step is to solve the differential Eqns. (4.20) and (4.21) from time t_{k-1} to time t_k (between observations). The result obtained represents a Gaussian distribution, or

$$p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1}) = N(\mathbf{x}(t_k)|\hat{\mathbf{x}}(t_k^-), \mathbf{P}(t_k^-)), \tag{4.22}$$

where the superscript $-$ indicates that those are values before incorporating an observation.

4.2.2 Update

Suppose we have an observation available at t_k. Starting from the observation model in Eq. (4.17), we have

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}(t_k), t_k) + \mathbf{v}_k, \tag{4.23}$$

and

$$p(\mathbf{y}_k|\mathbf{x}(t_k)) = N(\mathbf{y}_k|\mathbf{h}(\mathbf{x}(t_k), t_k), \mathbf{R}_k). \tag{4.24}$$

Since the EKF is a linear approximation, the expansion of the nonlinear function $\mathbf{h}(\mathbf{x}(t_k), t_k)$ using Taylor series is employed, thus

$$\mathbf{h}(\mathbf{x}(t_k), t_k) = \mathbf{h}(\hat{\mathbf{x}}(t_k^-), t_k) + \mathbf{H}(\hat{\mathbf{x}}(t_k)^-, t_k)(\mathbf{x}(t_k) - \hat{\mathbf{x}}(t_k^-)) + ... \tag{4.25}$$

where $\mathbf{H}(\hat{\mathbf{x}}(t_k^-), t_k)$ is the Jacobian matrix of $\mathbf{h}(\mathbf{x}(t_k), t_k)$, given by

$$\mathbf{H}(\hat{\mathbf{x}}(t_k^-), t_k) = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_m}{\partial x_1} & \frac{\partial h_m}{\partial x_2} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix} \Bigg|_{\mathbf{x}(t)=\hat{\mathbf{x}}(t_k^-)}$$

As before, the higher order terms are neglected. Thus Eq. (4.23) can be written as

$$\mathbf{y}_k = \mathbf{H}(\hat{\mathbf{x}}(t_k^-), t_k)\mathbf{x}(t_k) + \mathbf{h}(\hat{\mathbf{x}}(t_k^-), t_k) - \mathbf{H}(\hat{\mathbf{x}}(t_k^-), t_k)\hat{\mathbf{x}}(t_k^-) + \mathbf{v}_k \tag{4.26}$$

Define $\mathbf{u} := \mathbf{h}(\hat{\mathbf{x}}(t_k^-), t_k) - \mathbf{H}(\hat{\mathbf{x}}(t_k^-), t_k)\hat{\mathbf{x}}(t_k^-)$. Then we have

$$p(\mathbf{y}_k|\mathbf{x}(t_k)) = N(\mathbf{y}_k|\mathbf{H}\mathbf{x}(t_k) + \mathbf{u}, \mathbf{R}_k). \tag{4.27}$$

With the Gaussian distributions in Eq. (4.22) and Eq. (4.27) we can calculate the joint distribution using the property of a Gaussian distribution shown in Eq. (4.14), obtaining

$$\begin{aligned} p(\mathbf{x}(t_k), \mathbf{y}_k|\mathbf{y}_{1:k-1}) &= p(\mathbf{y}_k|\mathbf{x}(t_k))p(\mathbf{x}(t_k)|\mathbf{y}_{1:k-1}) \\ &= N(\mathbf{y}_k|\mathbf{H}\mathbf{x}(t_k) + \mathbf{u}, \mathbf{R}_k)N(\mathbf{x}(t_k)|\hat{\mathbf{x}}(t_k^-), \mathbf{P}(t_k^-)) \\ &= N\left(\begin{bmatrix} \mathbf{x}(t_k) \\ \mathbf{y}_k \end{bmatrix} \middle| \begin{bmatrix} \hat{\mathbf{x}}(t_k^-) \\ \mathbf{H}\hat{\mathbf{x}}(t_k^-) + \mathbf{u} \end{bmatrix}, \begin{bmatrix} \mathbf{P}(t_k^-) & \mathbf{P}(t_k^-)\mathbf{H}^T \\ \mathbf{H}\mathbf{P}(t_k^-) & \mathbf{H}\mathbf{P}(t_k^-)\mathbf{H}^T + \mathbf{R}_k \end{bmatrix} \right). \end{aligned} \tag{4.28}$$

From this result (similar structure to the in Eq. (4.15) and considering that

$\mathbf{h}(\hat{\mathbf{x}}(t_k^-), t_k) = \mathbf{H}\hat{\mathbf{x}}(t_k^-) + \mathbf{u}$), and using the results from Eqns. (4.16), $p(\mathbf{x}(t_k)|\mathbf{y}_k)$ is obtained as

$$p(\mathbf{x}(t_k)|\mathbf{y}_k) = N(\hat{\mathbf{x}}(t_k), \mathbf{P}(t_k)) \tag{4.29}$$

where

$$\hat{\mathbf{x}}(t_k) = \hat{\mathbf{x}}(t_k^-) + \mathbf{P}(t_k^-)\mathbf{H}^T[\mathbf{H}\mathbf{P}(t_k^-)\mathbf{H}^T + \mathbf{R}_k]^{-1}[y_k - \mathbf{h}(\hat{\mathbf{x}}(t_k^-), t_k)] \tag{4.30}$$

$$\begin{aligned}\mathbf{P}(t_k) &= \mathbf{P}(t_k^-) - \left[\mathbf{P}(t_k^-)\mathbf{H}^T[\mathbf{H}\mathbf{P}(t_k^-)\mathbf{H}^T + \mathbf{R}_k]^{-1}\right] \\ &\quad [\mathbf{H}\mathbf{P}(t_k^-)\mathbf{H}^T + \mathbf{R}_k]\left[\mathbf{P}(t_k^-)\mathbf{H}^T[\mathbf{H}\mathbf{P}(t_k^-)\mathbf{H}^T + \mathbf{R}_k]^{-1}\right]^T\end{aligned} \tag{4.31}$$

The EKF can be recognized in Eq. (4.30) and Eq. (4.31). The variance of $\mathbf{y}_k|\mathbf{y}_{k-1}$, $\mathbf{S}_k$ and the Kalman Gain $\mathbf{K}_k$ are given by

$$\mathbf{S}_k = \mathbf{H}\mathbf{P}(t_k^-)\mathbf{H}^T + \mathbf{R}_k \tag{4.32}$$

and

$$\mathbf{K}_k = \mathbf{P}(t_k^-)\mathbf{H}^T\mathbf{S}_k^{-1}. \tag{4.33}$$

4.3 Implementation

The implementation for the continuous-discrete EKF (CDEKF) has a general structure that follows from considering the Bayesian framework. Some elements in the implementation are used in other filter strategies, in this case the particle filter described in Chapter 5.

Since the CDEKF has a continuous time element, the prediction step requires

solution of two ordinary differential equations between observations, Eq. (4.20) and Eq. (4.21). These differential equations usually require numerical methods for integration even though, Eq. (4.21) is linear and could be solved by analytical methods but it has high dimensionality in this case. In this thesis, Runge Kutta of order 4/5 (RK4) is used since it uses a constant step size and has short computation time compared with other methods, and we can vary the step size between iterations.

A vital step in the implementation of Bayesian filters and state estimators is to obtain a state space representation of the system. There are infinite state space representations of a system [35]. In our case the state space form used was presented in Chapter 2. Adding noise $\mathbf{w}(t)$ to this state space form will render the model a stochastic differential equation [11]. For the CDEKF is also important to obtain an analytic representation of the Jacobian matrices which can also be checked using numerical methods.

Finally the observation model is needed. In the case of tracking objects in the atmosphere of a planet, radars have been the most widely used sensor when internal control and knowledge of the object is not at hand. Radars are complicated systems. Since the aim of this work is not focused on the measurement system but in the modeling of the spiraling target, a simplistic approach to a radar model is used [36], where the measurements obtained are range ρ, rate $\dot{\rho}$, elevation and azimuth, or

$$\begin{aligned} \mathbf{s} &= \mathbf{r} - \mathbf{r}_{r,n} \\ \rho &= \|\mathbf{s}\| \\ \dot{\rho} &= \frac{\dot{\mathbf{r}} \odot \mathbf{s}}{\|\mathbf{s}\|} \\ azi &= \arcsin \frac{(\mathbf{s} * \mathbf{k}) \cdot \mathbf{i}}{\|\mathbf{s} * \mathbf{k}\|} \\ elev &= \pi/2 - \arcsin \frac{\mathbf{s} \cdot \mathbf{k}}{\|\mathbf{s} * \mathbf{k}\|} \end{aligned} \tag{4.34}$$

where $\mathbf{r}_{r,n}$ is the position were the radar is located in the planet-centered, inertial reference frame and $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ are the orientation vectors in the local East-North-Up frame in which the measurements are defined. The quantities from Eq. (4.34) form the observations vector at t_k given by

$$\mathbf{h}(\mathbf{x}(t_k), t_k) = \left[\rho_k, azi_k, elev_k, \dot{\rho}_k\right] \tag{4.35}$$

The algorithm to implement the CDEKF is

1 set $\mathbf{P}(t) = \mathbf{P}_0, \hat{\mathbf{x}}(t) = \hat{\mathbf{x}}(t_0)$, $t_0 \leq t \leq t_T$, $1 \leq k \leq T$

2 from/for $t = t_0$ to t_T do

3 evaluate $\mathbf{F}$ with $\hat{\mathbf{x}}(t_{k-1})$

4 for t_{k-1} to t_k

5 make state prediction solving $\frac{\mathrm{d}}{dt}\hat{\mathbf{x}}(t) = f(\hat{\mathbf{x}}(t), t)$

6 get solution $\hat{\mathbf{x}}(t_k^-)$

7 solve $\frac{d\mathbf{P}}{dt} = \mathbf{F}\mathbf{P} + \mathbf{P}\mathbf{F}^T + \mathbf{Q}(t)$

8 get solution $\mathbf{P}(t_k^-)$

9 endfor

10 make measurement prediction $\hat{\mathbf{y}}_k = \mathbf{h}(\hat{\mathbf{x}}(t_k^-), t_k)$

11 evaluate $\mathbf{H}$ with $\hat{\mathbf{x}}(t_k^-)$

12 calculate $\mathbf{S}_k = \mathbf{H}\mathbf{P}(t_k^-)\mathbf{H}^T + \mathbf{R}_k$

13 calculate $\mathbf{K}_k = \mathbf{P}(t_k^-)\mathbf{H}^T\mathbf{S}_k^{-1}$

14 calculate estimate (mean) $\hat{\mathbf{x}}(t_k) = \hat{\mathbf{x}}(t_k^-) + \mathbf{K}_k[\mathbf{y}_k - \hat{\mathbf{y}}_k]$

15 calculate covariance $\mathbf{P}(t_k) = \mathbf{P}(t_k^-) - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^T$

16 store results $\hat{\mathbf{x}}(t_k)$ and $\mathbf{P}(t_k)$

17 endfor

The values of $\mathbf{P}_0$, $\hat{\mathbf{x}}_0$, $\mathbf{Q}(t)$ and $\mathbf{R}_k$ are important design parameters that determine how the filter is going to behave. The initial covariance $\mathbf{P_0}$ give us an area where the first estimate could be around the real values of the states. The initial estimate $\mathbf{x}(t_0)$ is an initial guess of the values of the states to be estimated. Its choice can be based on raw measurements or knowledge about the system. When performing Monte Carlo simulations the initial estimate is chosen randomly inside the set of values of the initial covariance for each run. The process noise $\mathbf{Q}(t)$ is determined taking in account how accurate we consider our model to be, and how much information is lost after the linear approximation. The sensor noise $\mathbf{R}_k$ in theory is defined by the tolerances presented by the sensor on its different measurements, established by the manufacturer of the sensor. It can be changed depending on the application and the behavior observed in the filter.

CHAPTER 5

Particle Filter

The concepts associated with the particle filter are presented. The techniques that form the basis of the filter include Monte Carlo methods, perfect sampling and importance sampling. The continuous-discrete particle filter is developed and key details of the implementation are presented.

5.1 Monte Carlo Methods

Monte Carlo methods are in reference to the Principality of Monaco since it was known as the capital of the "games of chance". The systematic development of Monte Carlo methods is dated to the 1940's, being born thanks to the work of the pioneers of computation, particle physics and hydrodynamics (probabilistic diffusion) working towards development of the atomic bomb [9]. Monte Carlo methods are numerical techniques to calculate probabilities and other related quantities using sequences of random numbers. For the case of one variable a general procedure would proceed as follows:

- Generate a sequence of random numbers, $r_1, r_2, r_3, \ldots, r_N$ uniformly distributed in the interval $[0, 1]$.
- Use this sequence to produce a new sequence, $x_1, x_2, x_3, \ldots, x_N$ distributed according to a given pdf of interest.

- Use the sequence of values x to estimate some property of a function $f(x)$. The values of x can be treated as simulated measurements and the probability of x taking values in an identified region can be estimated.

Formally a Monte Carlo calculation in our context is an integration. It is different from regular numerical methods since probability theory elements are used and the overall estimation error decreases with the square root of the number of samples N. There are many techniques inside the family of Monte Carlo methods (see [37] for example). but the set of techniques known as sequential Monte Carlo methods and sampling methods are the techniques that led to the development of the particle filter.

5.2 Exact Sampling

Assume we have N independent and identically distributed random samples (also named as particles), $\left\{\mathbf{x}(t_{0:k})^{(i)}, i = 1, 2, \ldots, N\right\}$ according to $p(\mathbf{x}(t_{0:k})|y_{1:k})$. An empirical estimate of the distribution can be given by

$$p^N(\mathbf{x}(t_{0:k})|y_{1:k}) = \frac{1}{N}\sum_{i=1}^{N}\delta_{\mathbf{x}(t_{0:k})^{(i)}} \tag{5.1}$$

where $\delta_{\mathbf{x}(t_{0:k})^{(i)}}$ denotes the Dirac function located in $\mathbf{x}(t_{0:k})^{(i)}$. Using standard probability theory, the expected value of a function $E_{p^N}[f]$ can be calculated.

$$E_{p^N}[f] = \int f(\mathbf{x}(t_{0:k}))p^N(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})d\mathbf{x}(t_{0:k}) = \frac{1}{N}\sum_{i=1}^{N} f(\mathbf{x}(t_{0:k})^{(i)}). \tag{5.2}$$

Since $p^N(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})$ is and estimate of $p(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})$, then Eq. (5.2) is a discrete estimate of the expected value of a function $E_p[f]$

$$\begin{aligned} \hat{E}^N[f] = E_{p^N}[f] &\simeq E_p[f] \\ \int f(\mathbf{x}(t_{0:k}))p^N(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})d\mathbf{x}(t_{0:k}) &\simeq \int f(\mathbf{x}(t_{0:k}))p(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})d\mathbf{x}(t_{0:k}). \end{aligned}$$

This estimate is unbiased and if the posterior variance $\sigma_f^2 < \infty$, then the variance of $\hat{E}^N(f)$ is equal to $\frac{\sigma_f^2}{N}$. From the strong law of numbers [9], we have

$$E_{p^N}[f] \xrightarrow[N\to\infty]{a.s} E_p[f], \tag{5.3}$$

where $\xrightarrow{a.s}$ denotes almost convergence. Also if $\sigma_f^2 < \infty$, then the central limit theorem holds and

$$\sqrt{N}\big[E_{p^N}[f] - E_p[f]\big] \Rightarrow N(0, \sigma_f^2) \tag{5.4}$$

where $\Rightarrow$ denotes convergence in distribution. One significant advantage over deterministic numerical integration methods is that the accuracy is independent of the dimension of the integrand [37]. It is extremely challenging to sample the posterior directly, especially in the case of a large dimensional state space. Due to this technical problem, different techniques are used in practice. There are several key sampling techniques [37]. The particle filter was developed using importance sampling.

5.3 Importance Sampling

The basic idea of importance sampling is to use a new density function called importance distribution (also referred as proposal distribution or instrumental distribution),

denoted as $\pi(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})$, and then to use weights to correct the fact that the sampling is made from the importance distribution instead of the target distribution $p(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})$. Consider the identity [38]

$$\begin{aligned} P(\mathbf{x} \in X) = \int_X p(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k}) d\mathbf{x}(t_{0:k}) &= \int_X \pi(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k}) \frac{p(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})}{\pi(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})} d\mathbf{x}(t_{0:k}) \\ &= \int_X \pi(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k}) w(\mathbf{x}(t_{0:k})) d\mathbf{x}(t_{0:k}) \end{aligned} \tag{5.5}$$

Then with N independent $\left\{\mathbf{x}(t_{0:k})^{(i)}, i = 1, 2, \ldots, N\right\}$ samples from $\pi(\mathbf{x}(t_{0:k})|y_{1:k})$ and computing the expected value of $f(\mathbf{x}(t_{0:k}))$, it follows that

$$\begin{aligned} E[f(\mathbf{x}(t_{0:k})] &= \int_X f(\mathbf{x}(t_{0:k}) \pi(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k}) w(\mathbf{x}(t_{0:k})) d\mathbf{x}(t_{0:k}) \\ &\simeq \frac{1}{N} \sum_{i=1}^{N} w(\mathbf{x}(t_{0:k})^{(i)}) f(\mathbf{x}(t_{0:k})^{(i)}) \\ \hat{E}^N[f] &= \frac{1}{N} \sum_{i=1}^{N} w_k^{(i)} f(\mathbf{x}(t_{0:k})^{(i)}) \end{aligned} \tag{5.6}$$

For some practical applications the quotient of the probability functions will usually not be normalized when forming the weight $w_k^{(i)}$. For this reason a normalization has to de employed. Hence we have

$$\hat{E}^N[f] = \sum_{i=1}^{N} w_k^{(i)} f(\mathbf{x}(t_{0:k})^{(i)}), \tag{5.7}$$

where,

$$w_k^{(i)} = \frac{w_k^{(i)}/N}{\sum_{j=1}^{N} w_k^{(j)}/N}. \tag{5.8}$$

The estimate in Eq. (5.7) presents the same characteristics shown at the end of exact

sampling section. In other words,

$$\sqrt{N}\big[E^N[f] - E[f]\big] \Rightarrow N(0, \sigma_f^2). \tag{5.4}$$

5.4 Particle Filter

The particle filter takes advantage of sampling, but not in a batch fashion. For state estimation, a sequential (recursive) representation is important. This is achieved by finding a relation with the general Bayes filter. The result is known as sequential importance sampling and constitutes a basic version of the particle filter.

5.4.1 Sequential Importance Sampling

The importance sampling method can be modified so that the values obtained at an instant t_{k-1} are used to obtain an estimate. This implies using the past simulated trajectories (samples) $\big\{\mathbf{x}(t_{0:k-1})^{(i)}, i = 1, 2, \ldots, N\big\}$ and importance weights $w_{0:k-1}^{(i)}$, which in turn implies the use of the Markov assumption and the conditionally independence of observations assumption. This means that the importance distribution $\pi(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})$ follows the Markov assumption and can be represented as

$$\pi(\mathbf{x}(t_{0:k})|y_{1:k}) = \pi(\mathbf{x}(t_k)|\mathbf{x}(t_{0:k-1}), \mathbf{y}_{1:k})\pi(\mathbf{x}(t_{0:k-1})|\mathbf{y}_{1:k-1}),$$

or iterating, we obtain

$$\pi(\mathbf{x}(t_{0:k})|y_{1:k}) = \prod_{i=1}^{k} \pi(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}), \mathbf{y}_{1:i})\pi(\mathbf{x}(t_0)). \tag{5.9}$$

This importance distribution can easily be used to evaluate sequentially in time the

importance weights. Then starting from the general expression for the importance weights and under the assumptions mentioned above, a sequential representation of w_k is obtained

$$\begin{aligned} w_k \equiv w(\mathbf{x}(t_{0:k})) &= \frac{\overbrace{p(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})}^{Eq.\ (3.5)}}{\pi(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})} \\ &\propto \frac{p(\mathbf{y}_{1:k}|\mathbf{x}(t_{0:k}))p(\mathbf{x}(t_{0:k}))}{\pi(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})} \\ &\propto \prod_{i=1}^{k} \frac{p(\mathbf{y}_{1:i}|\mathbf{x}(t_{0:i}))p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}))p(\mathbf{x}(t_0))}{\pi(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}),\mathbf{y}_{1:i})\pi(\mathbf{x}(t_0))}. \end{aligned} \tag{5.10}$$

Expanding for the k^{th} term we get a recursive representation for w_k as

$$\begin{aligned} w_k &\propto \prod_{i=1}^{k-1} \frac{p(\mathbf{y}_{1:i}|\mathbf{x}(t_{0:i}))p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1})p(\mathbf{x}(t_0))}{\pi(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}),\mathbf{y}_{1:i})\pi(\mathbf{x}(t_0))} \frac{p(\mathbf{y}_k|\mathbf{x}(t_k))p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}))}{\pi(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}),\mathbf{y}_{1:k})} \\ &\propto w_{k-1}\frac{p(\mathbf{y}_k|\mathbf{x}(t_k))p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}))}{\pi(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}),\mathbf{y}_{1:k})}. \end{aligned} \tag{5.11}$$

This expression is presented proportional because the denominator in Eq. (3.5) was not taken in account. In general the problem of implementing a filter with the sequential importance sampling paradigm consists of choosing an appropriate importance distribution $\pi(\mathbf{x}(t_{0:k})|\mathbf{y}_{1:k})$ and likelihood function $p(\mathbf{y}_k|\mathbf{x}(t_k))$, then sampling from an initial distribution $\pi(\mathbf{x}(t_0))$ and using the transition distribution $p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1})$ to evolve the samples, and finally finding the weights for this samples. The weights are calculated with the sampled form of Eq. (5.11) as

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{y}_k|\mathbf{x}(t_k)^{(i)})p(\mathbf{x}(t_k)^{(i)}|\mathbf{x}(t_{k-1})^{(i)})}{\pi(\mathbf{x}(t_k)^{(i)}|\mathbf{x}(t_{k-1})^{(i)},\mathbf{y}_{1:k})}. \tag{5.12}$$

If necessary the weights can be normalized as seen in equation Eq. (5.8). This depends on the choice of distributions.

5.4.2 Continuous-Discrete Particle Filter

The particle filter is an algorithmic implementation of sequential importance sampling. First, we must choose the importance distribution $\pi(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}), \mathbf{y}_k)$. For the case of the first particle filter, the choice was

$$\pi(\mathbf{x}(t_{0:k})|y_{1:k}) = p(\mathbf{x}(t_{0:k})) = \prod_{i=1}^{k} p(\mathbf{x}(t_i)|\mathbf{x}(t_{i-1}))p(\mathbf{x}(t_0))$$

which minimizes the variance of the importance weights [31]. This particle filter is known as the bootstrap filter or condensation algorithm and was first introduced by Gordon et al [24]. Then the recursive weight equation (Eq. (5.12)) becomes

$$w_k^{(i)} = w_{k-1}^{(i)} p(\mathbf{y}_k|\mathbf{x}(t_k)^{(i)}).$$

The evolution of the set of samples is made by the transition distribution $p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}))$ that represents the evolution of $\mathbf{x}(t)$. After drawing an initial set of particles from $p(\mathbf{x}(t_0))$ the continuous-discrete particle filter can be shown in the prediction/update fashion [34]. To propagate trajectories $\left\{\mathbf{x}(t)^{(i)} : t_{k-1} \leq t \leq t_k, i = 1, 2, \ldots, N\right\}$, we have

$$\frac{\mathrm{d}\mathbf{x}^{(i)}(t)}{\mathrm{d}t} = \mathbf{f}(\mathbf{x}(t)^{(i)}, \mathbf{u}(t), t) + \sigma(\mathbf{x}(t)^{(i)}, t)\beta(t)^{(i)}$$

with $\mathbf{x}(t_{k-1})^{(i)}$ being the initial condition. Then the samples $\mathbf{x}(t_k)^{(i)}$ obtained as the solution have the distribution $p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1}))$ and are directly related with the probability density function of the noise $\beta(t)$.

The updates are incorporated when measurements are available using the likelihood function and the set of samples obtained, to calculate the new importance weights, given by

$$w_k^{(i)} = w_{k-1}^{(i)} p(\mathbf{y}_k | \mathbf{x}(t_k)^{(i)}).$$

These weights can be used in the computation of the point estimates or to carry out a resampling process (discussed in next section). The output can be represented as the set $\left\{\mathbf{x}(t_k)^{(i)}, w_k^{(i)}\right\}\ i = 1, 2, \ldots, N$.

5.5 Resampling

Particle filters can display degeneracy, which manifests itself after many iterations where one of the samples will have a relevant weight, while the remaining will be insignificant [12]. Degeneracy is a problem that cannot be completely avoided due to the characteristics of the algorithm that the variance of the weights increase with time [38]. This degeneracy can be decreased in two ways. First, by choosing adequate importance distribution (this changes the usual form of the bootstrap filter) and second, through resampling techniques. Resampling consists in removing the samples with insignificant weight, while maintaining and reproducing the samples that are important for the process (have a high importance weight), and then assigning the same weight to all the surviving samples. Formally it consists on retracting from a weighted representation of the posterior Eq. (5.7) to a standard sampled representation Eq. (5.1).

There are many resampling techniques and more could be developed using mathematical techniques or heuristics, but there is a set of well established techniques that can be readily used [12], [37], [39], [38]. In this case several techniques were tested, and the residual resampling was chosen based on the results from [39] which shows its superiority

over other techniques for a sample size $N = 1000$.

Residual resampling consists on allocating $n_i^{\cdot} = \lfloor N w_i \rfloor$ copies of particle $\mathbf{x}(t_k)^{(i)}$ to the new distribution. Additionally, resample $m = N \sum n_i^{\cdot}$ particles from $\left\{\mathbf{x}(t_{0:k})^{(i)}, i = 1, 2, \ldots, N\right\}$ by making $n_i^{\cdot\cdot}$ copies of particle $\mathbf{x}(t_k)^{(i)}$ where the probability for selecting $\mathbf{x}(t_k)^{(i)}$ is proportional to $w_i^{\cdot} = N w_i n_i^{\cdot}$ [39].

5.6 Implementation

A general implementation of the algorithm for the continuous-discrete particle filter is relatively simple, but has several practical issues that have to be addressed to avoid filter divergence. In the case of this work, since the values of position, velocity and acceleration are measured in the planet-centered inertial reference frame, the values of the state are large and numerical problems can arise, so appropriate scaling is necessary.

5.6.1 Likelihood Function

An important step in the implementation is the choice of the likelihood function. It can be chosen as any probability density function. Since the comparison is being made with a Gaussian approximation (EKF), the noise (process and measurement) is then chosen to have this same nature. From this follows that the likelihood function should have itself a Gaussian distribution.

When choosing the likelihood function for a radar sensor that gives different measurement types (e.g., range and rate, or azimuth and elevation) we could assume the independence of the measurements. Choosing a Gaussian representation for each we can obtain the final likelihood function (e.g., [40]). This assumption does not use the

information contained between the particles that come from the prediction step and can easily cause divergence of the particle filter. To account for the state prediction,we have a multivariate Gaussian distribution with a covariance that gives a relationship between the measurement and sample covariance given by

$$p(\mathbf{y}_k|\mathbf{x}(t_k)) = N(\mathbf{y}_k|\mathbf{h}(\mathbf{x}(t_k), t_k), \mathbf{HP}(t_k^-)\mathbf{H}^T + \mathbf{R}_k), \tag{5.13}$$

where $\mathbf{H}$ is the Jacobian of the measurements, $\mathbf{P}(t_k^-)$ is the covariance of the samples (calculated empirically) and $\mathbf{R}_k$ is the measurement covariance.

5.6.2 Algorithm

The particle filter approximates the posterior pdf using sets of state samples (particles):

$$\mathbf{X}_k \quad = \quad \{\mathbf{x}(t_k)^{(i)}\}_{i=1,\ldots,N} \tag{5.14}$$

The set $\mathbf{X}_k$ consists of N particles $\mathbf{x}(t_k)^{(i)}$, for some large number of N (e.g, $N = 1000$). Together, these particles approximate the posterior $p(\mathbf{x}(t_k)|\mathbf{y}_k)$. The set $\mathbf{X}_k$ is calculated recursively.

Initially, at the instant t_0, the particles $\mathbf{x}(t_0)^{(i)}$ are generated from the initial state distribution $p(\mathbf{x}(t_0))$. The k-th particle set $\mathbf{X}_k$ is then calculated recursively from $\mathbf{X}_{k-1}$ as follows:

1 set $\mathbf{X}_k = \mathbf{X}_k^{\text{aux}} = \emptyset$

2 for $j = 1$ to N do

3 pick the j-th sample $\mathbf{x}(t_{k-1})^{(j)} \in \mathbf{X}_{k-1}$

4 draw $\mathbf{x}(t_k)^{(j)} \sim p(\mathbf{x}(t_k)|\mathbf{x}(t_{k-1})^{(j)})$

5 set $w_k^{(j)} = p(\mathbf{y}_k|\mathbf{x}(t_k)^{(j)})$

5.1 normalize $w_k^{(j)}$

6 add $\langle \mathbf{x}(t_k)^{(j)}, w_k^{(j)} \rangle$ to $\mathbf{X}_k^{\text{aux}}$

7 endfor

8 for $i = 1$ to N do

9 draw $\mathbf{x}(t_k)^{(i)}$ from $\mathbf{X}_k^{\text{aux}}$ with probability proportional to $w_k^{(i)}$

10 add $\mathbf{x}(t_k)^{(i)}$ to $\mathbf{X}_k$

11 endfor

Lines 2 through 7 generates a new set of particles, with line 4 being the prediction step (solving an SDE). Lines 8 through 10 apply residual resampling (or a chosen resampling technique). Finally, after obtaining the set of particles from resampling, line 10, the important statistics can be calculated.

CHAPTER 6

Results and Analysis

6.1 Model Simulation Results

The class of trajectories that can be generated with the vehicle mathematical model are illustrated in this section. This model can be applied to any planet as long as a gravity and atmospheric models are available. Here we simulated trajectories on Mars and Earth.

Trajectories are created for a typical reentry vehicle whose tracking begins 50 km east of the origin of the local frame at an altitude of 100 km. The vehicle moves westward with an initial velocity of 1500 m/sec, oriented downward 10° relative to the local horizontal. The azimuth is chosen to be 180° so the trajectory follows a path parallel to the north direction and the different changes on the trajectory can be readily observed. The vehicle itself has the following properties:

$$\beta_m = 4000 \text{ Kg/m}^2 \qquad C_{D_0} = 0.033 \qquad (C_L/C_D)_{max} = 1.5 \qquad S = 1 \text{ m}$$

Several trajectories were generated using a periodic function of varying frequencies for $\dot{\varphi}$ and different values of λ. The values used represent plausible values for $\dot{\varphi}$ and λ. For each scenario, the trajectory torsion, the magnitude of the velocity and the lift are shown.

The two sets of numerical experiments (Earth and Mars) shown in Figures 6.1–6.8, are summarized in Table 6.1. In Figure 6.9, the final moments of the spiral depicted in

Figure 6.8 are magnified to more clearly illustrate the trajectory.

Table 6.1: Numerical Investigations of Spiraling Motion (for Mars and Earth)

$\dot{\varphi}$ (sec/turn)	λ
2π	0
20	0.6
2π	1
20	2

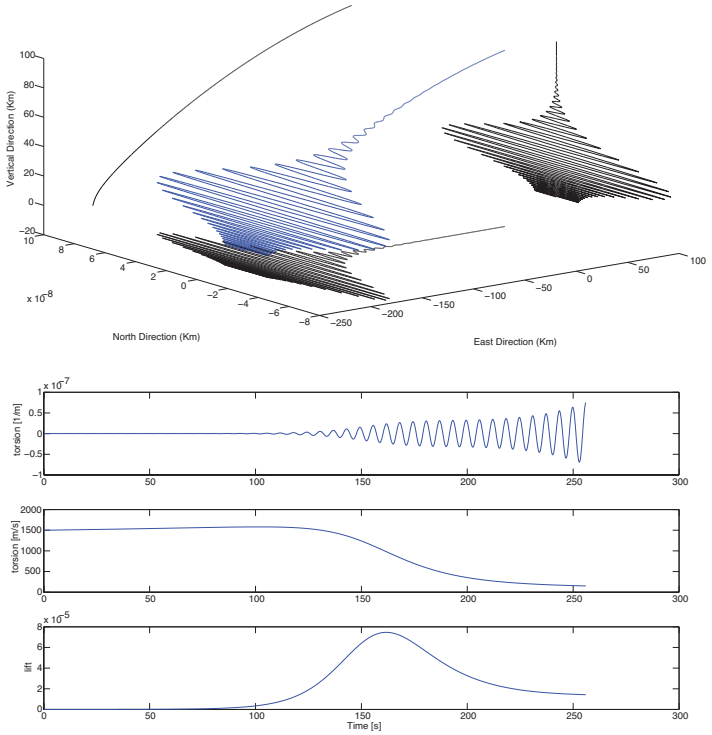

Figure 6.1: Trajectory for $\dot{\varphi} = 2\pi$ sec/turn and $\lambda = 0$ on Mars

6.1.1 Spiraling Change

It was observed distinct behavior for certain vehicle parameters as it manifested itself in the torsion computations. In particular, the torsion exhibited a "boundary" where the spiraling motion before and after had very different personalities. The torsion is a measure of the motion out of the osculating plane. A zero torsion means that the target trajectory

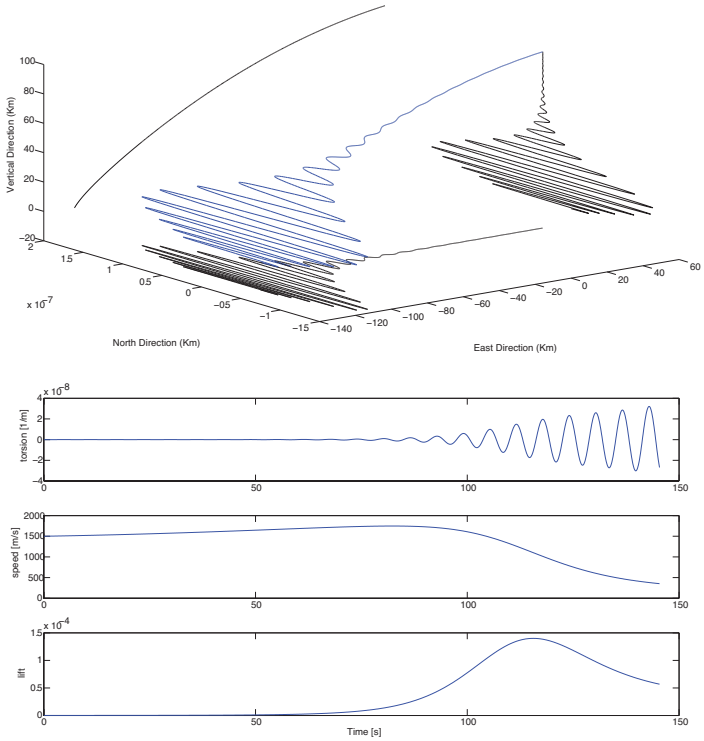

Figure 6.2: Trajectory for $\dot{\varphi} = 2\pi$ sec/turn and $\lambda = 0$ on Earth

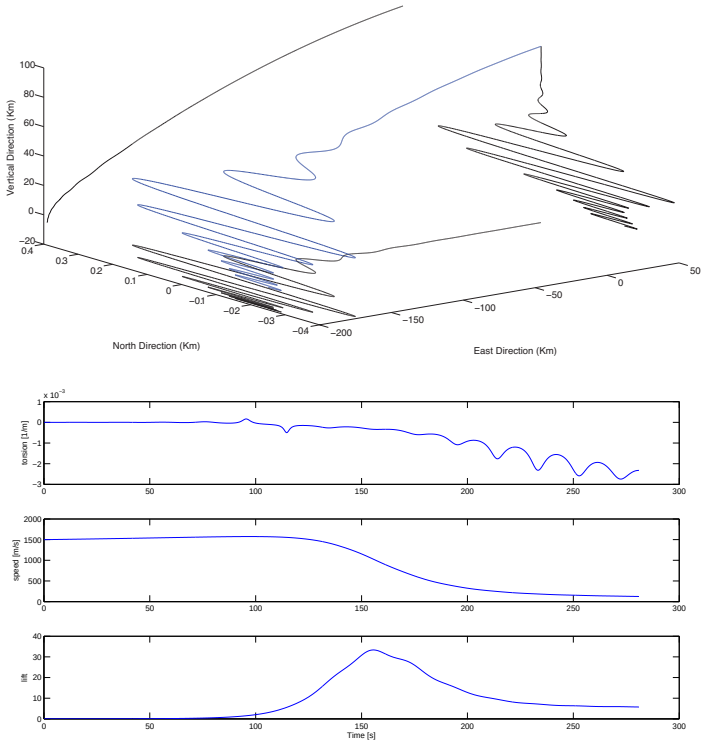

Figure 6.3: Trajectory for $\dot{\varphi} = 20$ sec/turn and $\lambda = 0.6$ on Mars

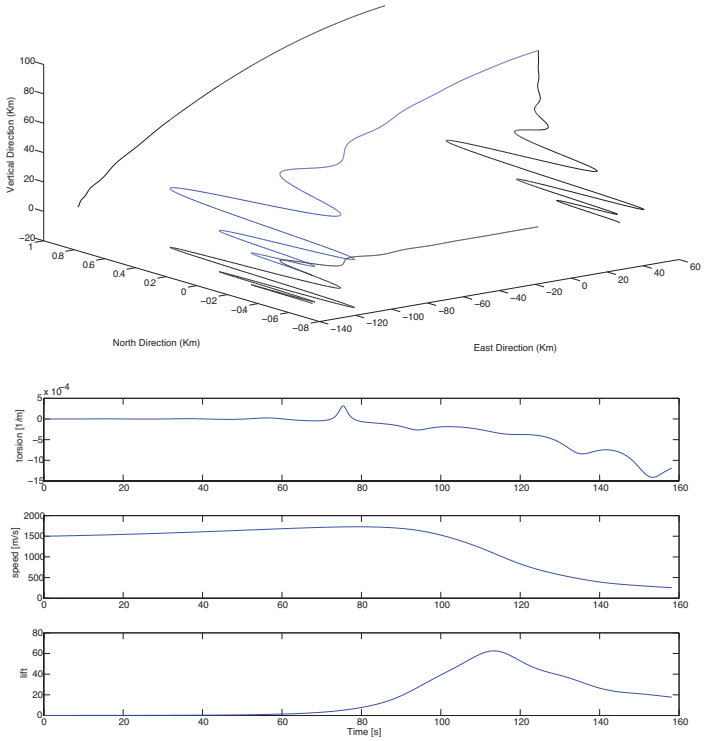

Figure 6.4: Trajectory for $\dot{\varphi} = 20$ sec/turn and $\lambda = 0.6$ on Earth

remains in a plane, known as the maneuver plane. The larger the magnitude of the torsion, the larger the displacement of the osculating plane. Note that the torsion can be a negative number, which is an indication of the direction of the spiraling relative to the osculating plane.

But what is then the relationship between the torsion and the actual spiraling as seen on the trajectory? It was observed that an out-of-plane motion (torsion different that zero) does not mean spiraling motion (see Fig. 6.10), but torsion has to be different than zero for this phenomenon to occur. The spiraling motion is actually related with the abrupt change of sign, "boundary", in the torsion. It was observed that this boundary is produced when several input parameters are different than zero: $\dot{\varphi}$ (rotation of the vehicle) and λ (related whit the generation of lift, and thus with the atmosphere and gravity). Significant spiraling rotation requires $\dot{\varphi} \neq 0$ and a significant value of lift generation.

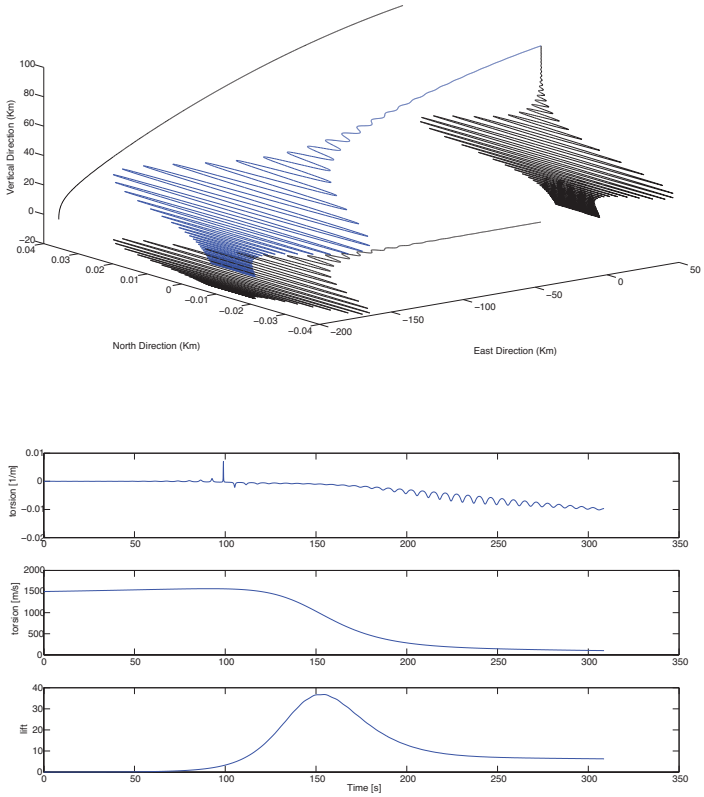

Figure 6.5: Trajectory for $\dot{\varphi} = 2\pi$ sec/turn and $\lambda = 1$ on Mars

6.2 Estimation Results

6.2.1 Simulated Trajectory For Estimation

To compare the two EKF and particle filter, it is important to choose a trajectory possessing characteristics important for the model (out of plane motion and spiraling) and characteristics that can be found in typical reentry trajectories. The general characteristics and parameters of the chosen trajectory are found in Table 6.2. These parameters produce a realistic trajectory where, without any control action, a change in torsion can be observed. Measurements are taken every $0.1s$ by a radar placed on the surface of the planet, and denoted by an asterisk in Figure 6.11. The origin of the local frame is placed at the radar location.

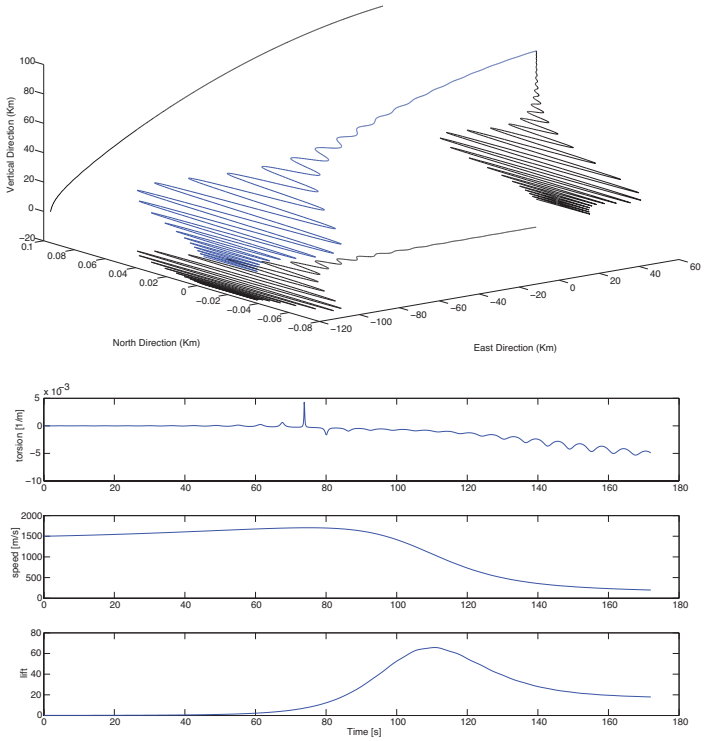

Figure 6.6: Trajectory for $\dot{\varphi} = 2\pi$ sec/turn and $\lambda = 1$ on Earth

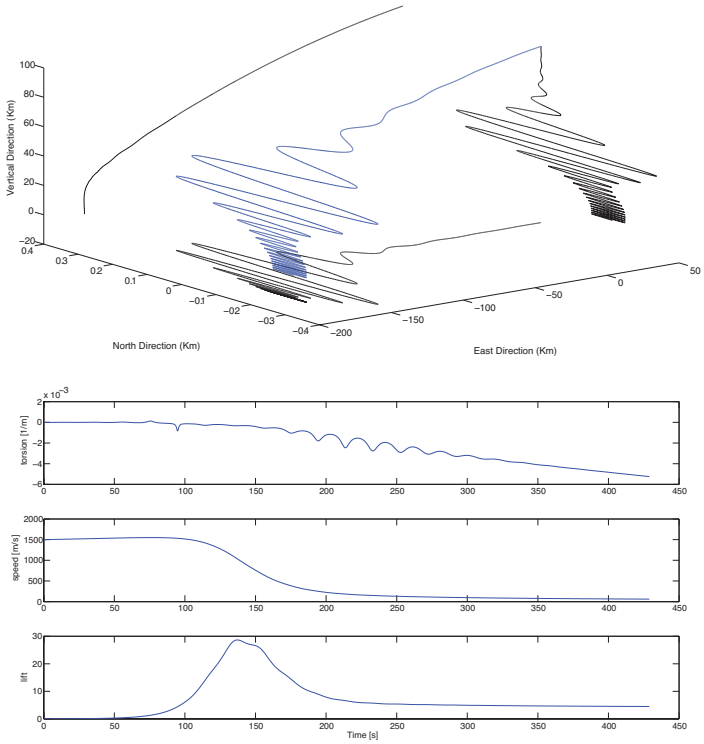

Figure 6.7: Trajectory for $\dot{\varphi} = 20$ sec/turn and $\lambda = 2$ on Mars

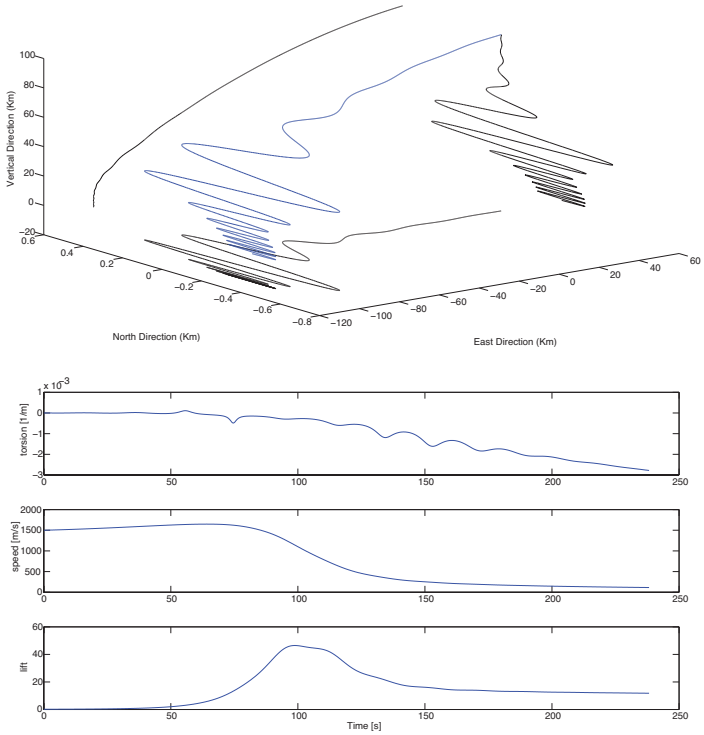

Figure 6.8: Trajectory for $\dot{\varphi} = 20$ sec/turn and $\lambda = 2$ on Earth

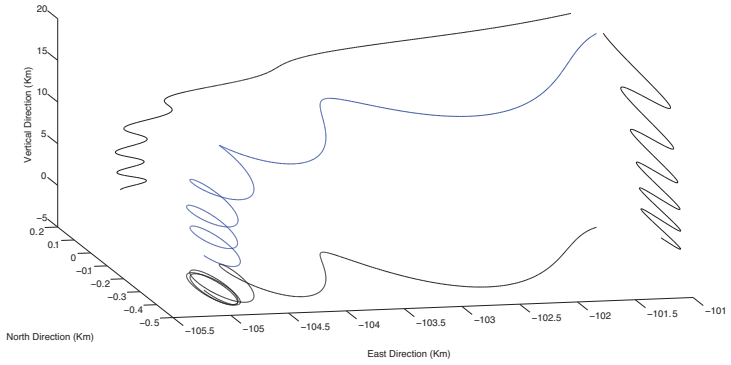

Figure 6.9: Detailed end of trajectory for $\dot{\varphi} = 20$ sec/turn and $\lambda = 2$ on Earth

The torsion and spiraling frequency for this trajectory were calculated and are displayed in Figures 6.12 and 6.13.

6.2.2 Filter Parameters

One of the most important and complicated steps in applied filtering is the tuning process of finding the parameters that make the filter function as desired. To compare (in an

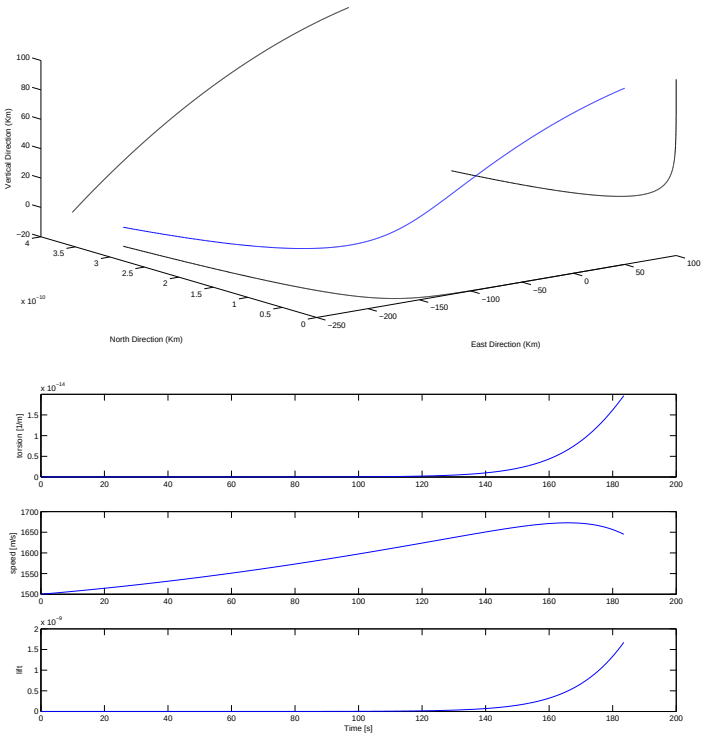

Figure 6.10: Trajectory for $\dot{\varphi} = 0$ sec/turn and $\lambda = 0$ on Mars

Table 6.2: Trajectory and vehicle parameters (speed and position in local frame)

Parameter	Value	units
$azimuth_0$	30	degrees
$elev_0$	-10	degrees
$speed_0$	5000	m/s
East	0	km
North	50	km
Altitude	100	km
λ_0	2	–
φ_0	0	rad
β_m	4000	kg/m^2
S	2	m
$\dot{\varphi}$	2π	sec/turn

objective way) the results given by the estimation techniques,it is essential to use the same values for the measurement noise, since the same sensor is used for both implementations. Dissimilar values for the process noise can be used, since, even though the same model is used, different kind of approximations are made in both filtering techniques.

Other important parameters are the time step used in the solution of the differential

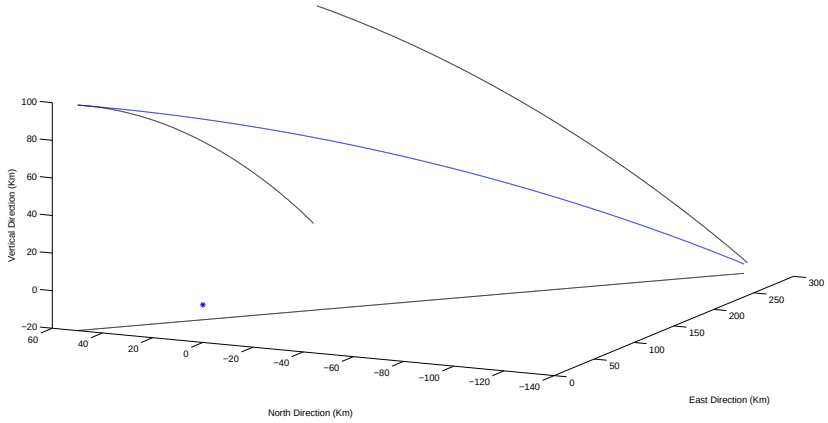

Figure 6.11: Chosen trajectory with radar (*), in local frame

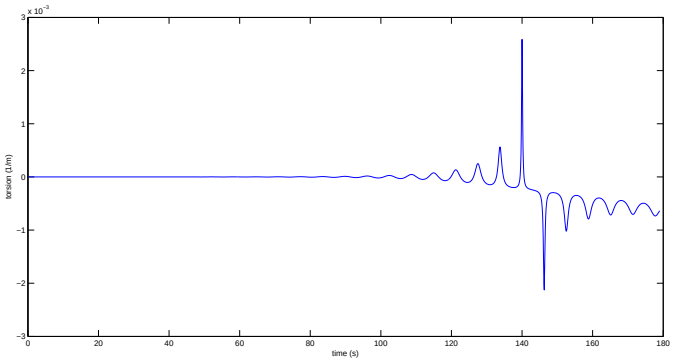

Figure 6.12: Torsion of the trajectory

equations. This affects the performance of the filters, since it determines the length of the prediction step. In this case, $\delta t = 0.03s$ for the EKF and $\delta t = 0.04s$ for the particle filter. The particle filter was made with N = 1000.

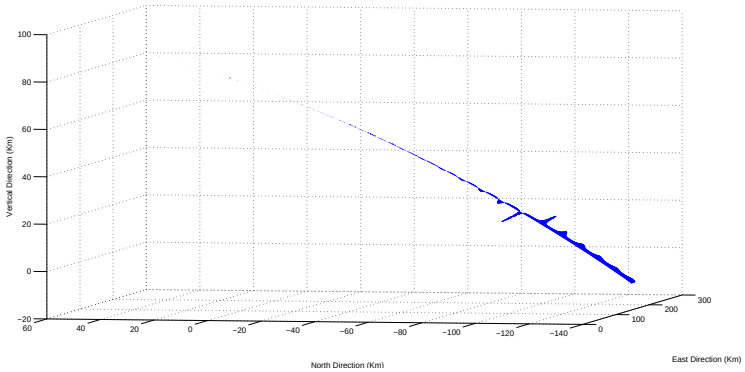

Figure 6.13: Change in spiraling frequency through the trajectory

Table 6.3: Parameter for the EKF

Parameter	Value
$\mathbf{Q}$	$diag([0\ 0\ 0\ 0\ 0\ 0\ 0.1^2\ 0.1^2\ 0.1^2\ 0.01^2\ 0.01^2])$
$\mathbf{R}$	$diag([10^2\ 0.001^2\ 0.001^2\ 10^2])$
$\mathbf{P}_0$	$diag([400^2\ 400^2\ 400^2\ 20^2\ 20^2\ 20^2\ 0.5^2\ 0.5^2\ 0.5^2\ 0.5^2\ 2\pi^2])$

Table 6.4: Parameter for the particle filter

Parameter	Value
$\mathbf{Q}$	$diag([0\ 0\ 0\ 0\ 0\ 0\ 0.01^2\ 0.01^2\ 0.01^2\ 0.01^2\ 0.01^2])$
$\mathbf{R}$	$diag([10^2\ 0.001^2\ 0.001^2\ 10^2])$
$\mathbf{P}_0$	$diag([400^2\ 400^2\ 400^2\ 20^2\ 20^2\ 20^2\ 0.5^2\ 0.5^2\ 0.5^2\ 0.5^2\ \pi^2])$

6.2.3 State Estimation Error

We have results showing the behavior of the difference between the truth (simulated trajectory) and the states estimated. Figures 6.14 to 6.21 show the error estimation for the states and the square root of their respective state estimation error covariance, after 30 Monte Carlo simulation runs.

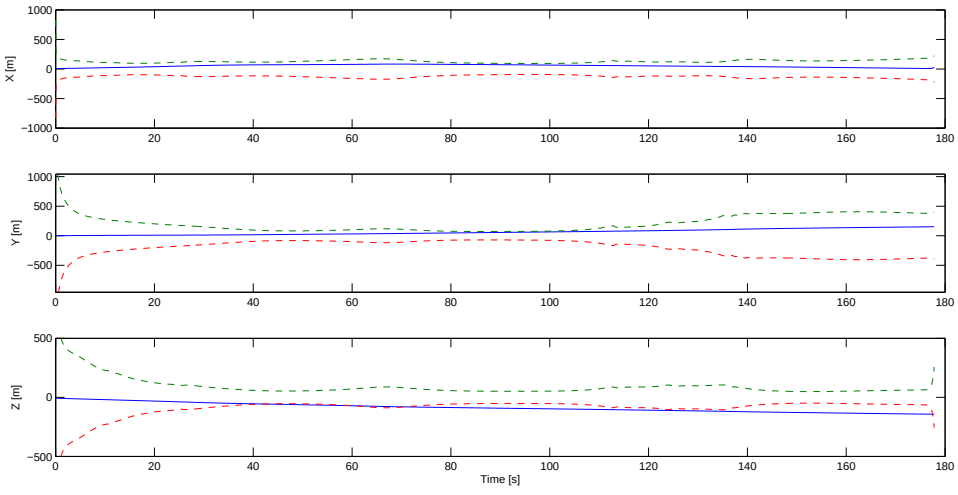

Figure 6.14: Position estimation error for the EKF

It is observed in some of the figures (see for example Fig. 6.20) that the covariance seems to grow again at the end of the trajectory. This is not due to divergence of the filter; it is due to the nature of the measurements, since the measured range has a similar behavior

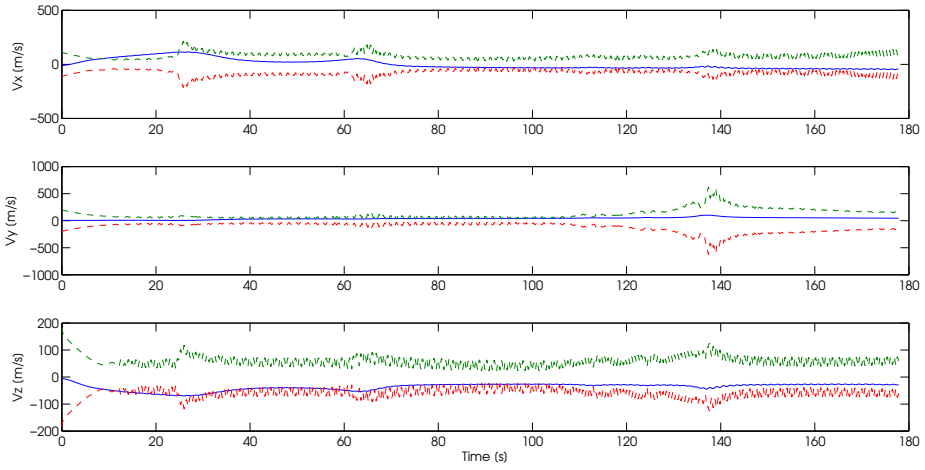

Figure 6.15: Velocity estimation error for the EKF

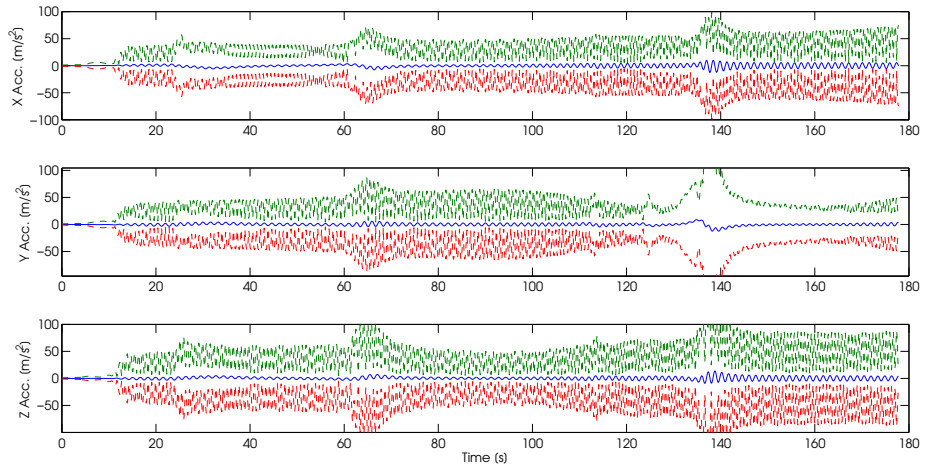

Figure 6.16: Acceleration estimation error for the EKF

(see Fig. 6.22).

6.2.4 Spiraling Estimation Error

The calculation of the spiraling is done through the calculation of the estimated torsion. Its error is a combination of the state estimation errors. The torsion is a small number, so in percentage the error is larger than other larger variables. The spiraling frequency error is obtained by scaling the torsion estimation errors with the velocity

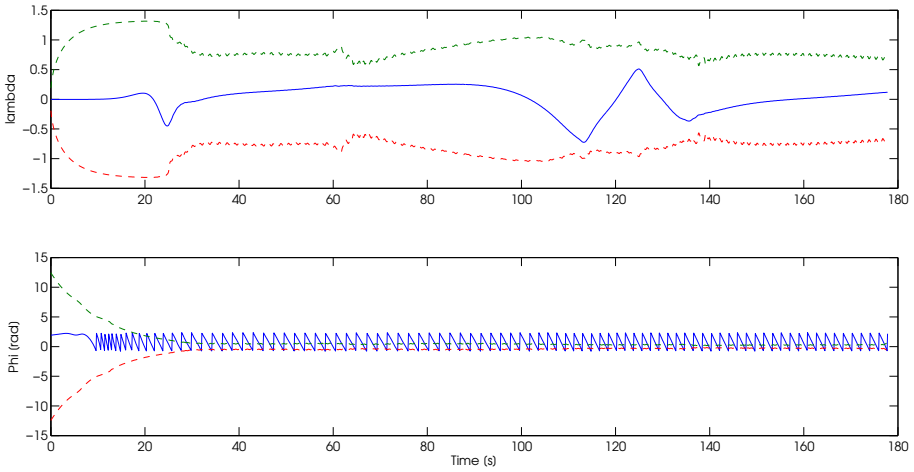

Figure 6.17: Parameters estimation error for the EKF

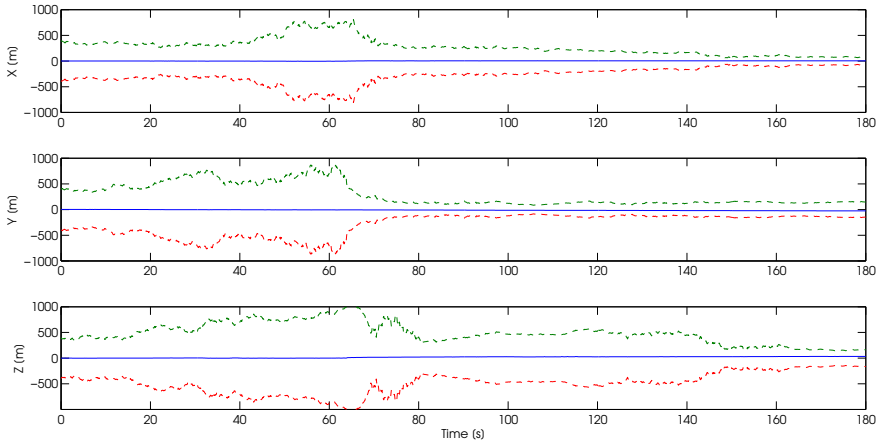

Figure 6.18: Position estimation error for the PF

estimation errors.

6.3 General analysis of estimation error

The successful implementation and errors obtained show that the coarse and highly nonlinear model can actually be used for estimation of the state and spiraling of a reentry vehicle, both techniques showing satisfactory results using just one radar. In general, even though in theory the covariance in the particle filter is suppose to increase with time (see Chapter 5), with a careful implementation a proper behavior of the covariance can be

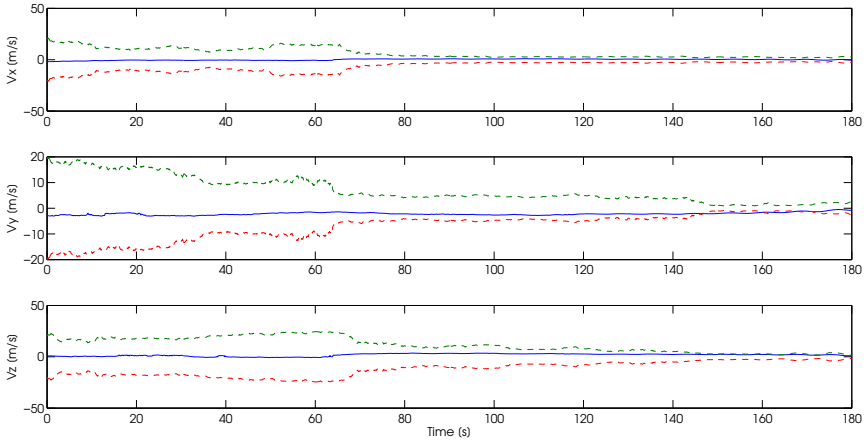

Figure 6.19: Velocity estimation error for the PF

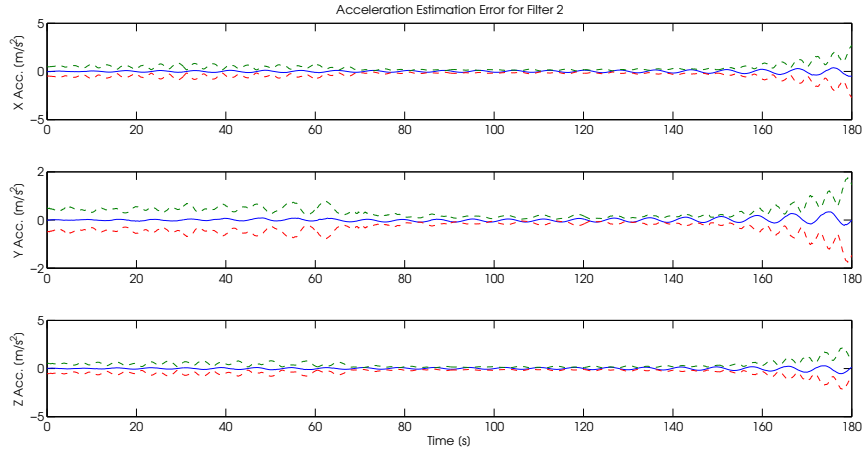

Figure 6.20: Acceleration estimation error for the PF

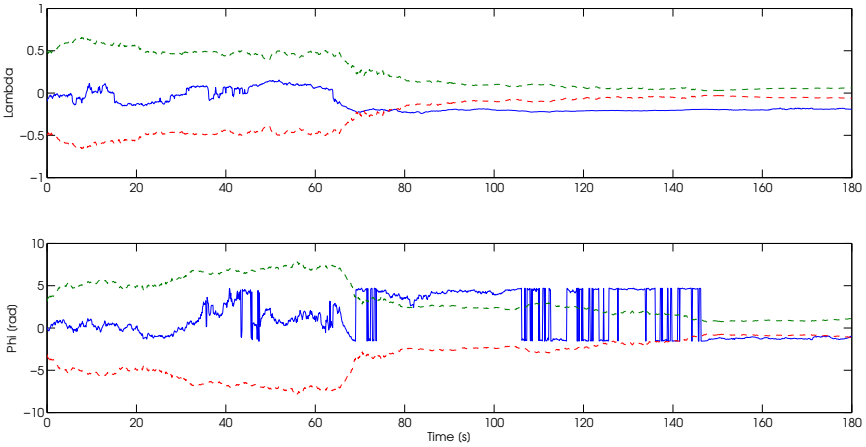

Figure 6.21: Parameters estimation error for the PF

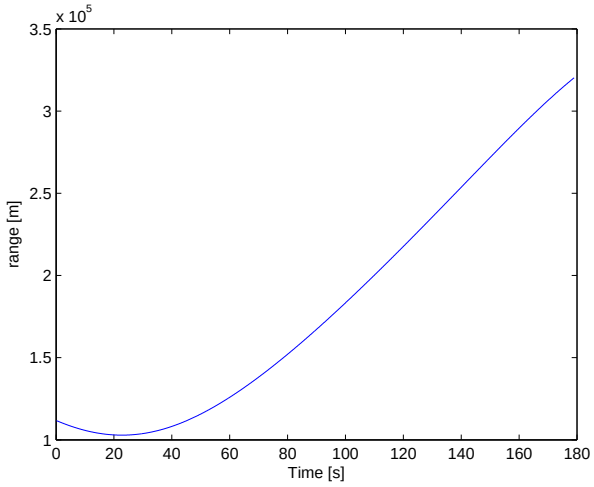

Figure 6.22: The range measurements increase with time (local frame)

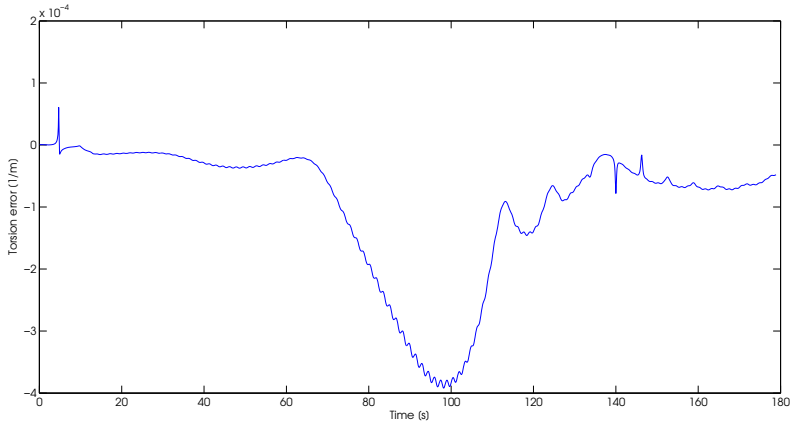

Figure 6.23: Torsion estimation error for the EKF

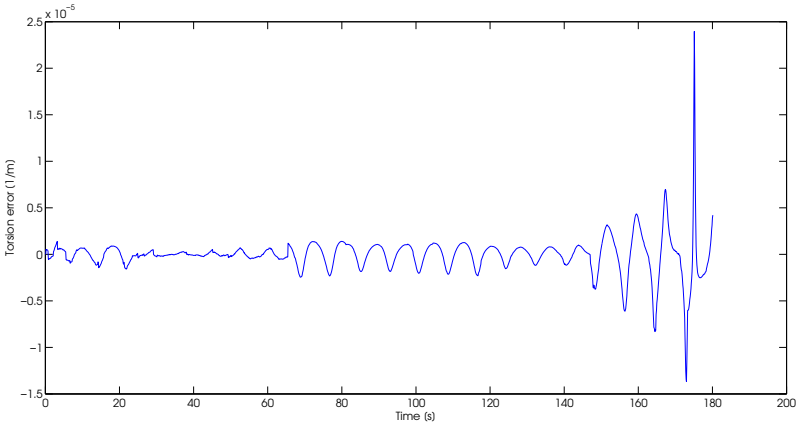

Figure 6.24: Torsion estimation error for the particle filter

obtained. The particle filter has analytical advantages, due to the sampled representation of the states at each instant of time. An empirical probability density function of the different variables can be obtained, for instance see Fig. 6.25, and it can be seen that it actually shows a non-Gaussian distribution, even though there was Gaussian assumption for the measurement and process noises.

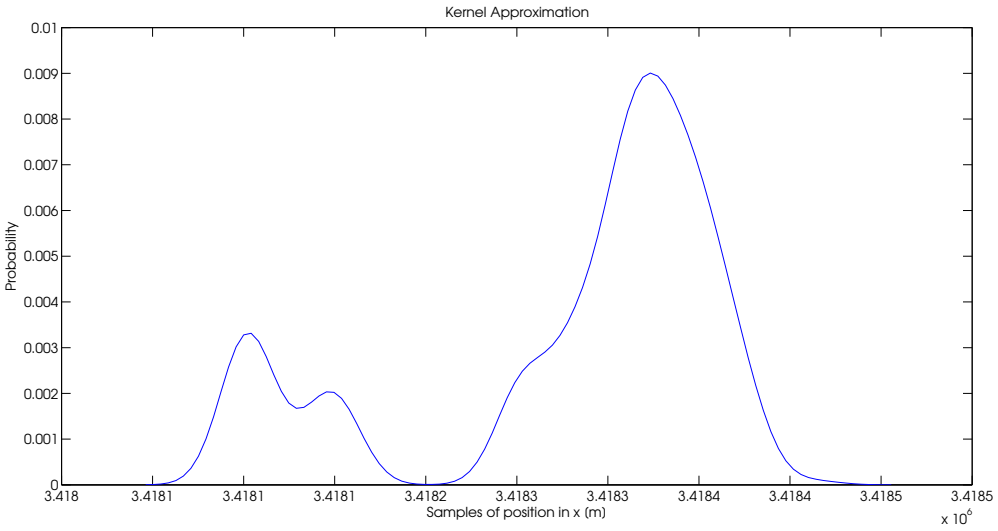

Figure 6.25: Probability density function for position in X at time = 167.6 s

In particular, it can be said that the particle filter outperformed the extended Kalman filter in the estimation of the parameters λ and φ. This can be seen comparing Figures 6.17 and 6.21, and is important when identifying a target. A better performance is obtained when estimating acceleration, Fig. 6.20, reflected too in a better estimation of the torsion from which it follows that we obtain a better estimation of the spiraling, in general.

CHAPTER 7

Conclusions

7.1 Conclusions

The main purpose of this work was to develop a physics-based model to be used for estimation of the state and spiraling of a reentry vehicle in the atmosphere. The types of reentry vehicles investigated are limited to axisymmetric vehicles. The estimation process was based on Bayesian estimation techniques, in particular, the popular extended Kalman filter and the newer technique known as the particle filter.

Concepts from differential geometry were used to analyze the trajectory behavior, using torsion as a measure of the spiraling of the trajectory. The torsion can also be scaled by velocity to obtain the spiraling frequency.

The use of the Bayesian framework is useful as a common foundation where the main technique is born naturally out of simple concepts of dynamical systems and probability theory, and then particular techniques are obtain by the application of key assumptions. This actually shows the relationship between the existing techniques and the possibilities of expansion and creation of new techniques.

The implementation of the EKF and the particle filter showed satisfactory results and can actually be used for the estimation of the states and torsion of the trajectory. It was shown that the particle filter can outperform the EKF, but for real time applications

parallelization will be needed for the particle filter to reduce the execution time.

7.2 Future Work

There are three particular lines of action where to extend and/or improve the results from this work. First, performing tests of an actual vehicle to validate the characteristics of a reentry trajectory model with high degree of confidence. A simplified version of the model was already implemented in a real system for aircraft tracking by Bishop [6]. Secondly, the model can be extended even more using other physical principles, adding more degrees of freedom to the vehicle movement, changing the treatment as a point to that of a rigid body. Lastly, a relationship between the particle filter and the technique of bank of expert EKFs [5] could be established.

APPENDIX A

Derivation of Jacobian Matrices

A.1 State Propagation Partials

The derivation of the state propagation partials for this model can be found in the Appendix of [5], but it doesn't include the parameters λ and φ. The partials here are just an extension of the ones shown in that work, expanding the size of the Jacobian matrix to include the parameters. Partials of the gravity and atmospheric models are not included because they are application dependant under the assumptions made about the planet. The derivation mentioned before is reproduced here to account for the different notation and additions to the Jacobian Matrix.

First, let recall that the planetary rotation vector is

$$\mathbf{\Omega} = [0 \;\; 0 \;\; \Omega]^T.$$

Then defining the relative state vector $\mathbf{X}_r$ we have

$$\begin{aligned} \dot{\mathbf{r}}_r &= \dot{\mathbf{r}} - \mathbf{\Omega} * \mathbf{r} \\ \ddot{\mathbf{r}}_r &= \mathbf{a} + \mathbf{g} - \mathbf{\Omega} * \dot{\mathbf{r}} \\ \mathbf{X}_r &= [\dot{\mathbf{r}}_r \;\; \ddot{\mathbf{r}}_r]^T \end{aligned} \tag{A.1}$$

with

$$\mathbf{e}_1^w = \frac{\dot{\mathbf{r}}_r}{r_r} \ , \quad \mathbf{e}_2^w = -\frac{\frac{\dot{\mathbf{r}}_r}{\dot{r}_r} * \mathbf{u}_z}{\left\| \frac{\dot{\mathbf{r}}_r}{\dot{r}_r} * \mathbf{u}_z \right\|} \ , \quad \text{and} \quad \mathbf{e}_3^w = \mathbf{e}_1^w * \mathbf{e}_2^w$$

We have the equation of motion

$$\begin{aligned} \ddot{\mathbf{r}} &= \mathbf{a} + \mathbf{g} \\ \dddot{\mathbf{r}} &= [\omega * \mathbf{a}] + [\dot{\varphi}\mathbf{e}_1^w * \mathbf{a}] - \left[\dot{D}\mathbf{e}_1^w\right] + \dot{L}\left[-\mathbf{e}_2{}^w \sin\varphi + \mathbf{e}_3{}^w \cos\varphi\right] + \dot{\mathbf{g}} \\ \mathbf{X} &= [\dot{\mathbf{r}} \ \ddot{\mathbf{r}}]^T \end{aligned} \tag{A.2}$$

with

$$\boldsymbol{\omega}^w = \frac{\dot{\mathbf{r}}_r * \ddot{\mathbf{r}}_r}{\dot{r}_r^2} \ ,$$

The first step is the derivation of the partials for the relative states $\mathbf{X}_r$:

$$\begin{aligned} \frac{\partial \mathbf{X}_r}{\partial \mathbf{X}} &= \begin{bmatrix} \mathbf{S}(\boldsymbol{\Omega}) & \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{S}(\boldsymbol{\Omega}) & \mathbf{I}_{3\times3} \end{bmatrix} \\ \frac{\mathrm{d}\mathbf{X}_r}{\mathrm{d}\mathbf{g}} &= [\mathbf{0}_{3\times3} \ \mathbf{I}_{3\times3}] \\ \frac{\mathrm{d}\mathbf{X}_r}{\mathrm{d}\mathbf{X}} &= \frac{\partial \mathbf{X}_r}{\partial \mathbf{X}} + \frac{\mathrm{d}\mathbf{X}_r}{\mathrm{d}\mathbf{g}} \frac{\mathrm{d}\mathbf{g}}{\mathrm{d}\mathbf{X}}, \end{aligned} \tag{A.3}$$

where $\mathbf{S}()$ designates the skew symmetric representing the vector cross product in the sense that

$$\mathbf{a} * \mathbf{b} = \mathbf{S}(\mathbf{a})\mathbf{b} \ \forall \ \mathbf{a}, \mathbf{b} \in \mathbb{R}^3.$$

We can also compute the partials of each of the wind frame vectors (omitting the

superscript w for easier reading)

$$
\begin{aligned}
\frac{\mathrm{d}\mathbf{e}_1}{\mathrm{d}\mathbf{X}_r} &= \frac{1}{\|\dot{\mathbf{r}}_r\|^3}\begin{bmatrix} \|\dot{\mathbf{r}}_r\|^2 - \dot{r}_{r_x}^2, & -\dot{r}_{r_x}\dot{r}_{r_y}, & -\dot{r}_{r_x}\dot{r}_{r_z} & \\ -\dot{r}_{r_x}\dot{r}_{r_y}, & \|\dot{\mathbf{r}}_r\|^2 - \dot{r}_{r_y}^2, & -\dot{r}_{r_y}\dot{r}_{r_z} & \mathbf{0}_{3\times 3} \\ -\dot{r}_{r_x}\dot{r}_{r_z}, & -\dot{r}_{r_y}\dot{r}_{r_z}, & \|\dot{\mathbf{r}}_r\|^2 - \dot{r}_{r_z}^2 & \end{bmatrix} \\
\frac{\partial \mathbf{e}_2}{\partial \mathbf{e}_1} &= -\left(-\frac{\mathbf{S}(\mathbf{r})}{\|\mathbf{e}_1 * \mathbf{r}\|} + 2\frac{(\mathbf{e}_1 * \mathbf{r})(\mathbf{S}(\mathbf{r})(\mathbf{e}_1 * \mathbf{r}))^T}{\|\mathbf{e}_1 * \mathbf{r}\|^3}\right) \\
\frac{\partial \mathbf{e}_3}{\partial \mathbf{e}_1} &= \mathbf{S}(\mathbf{e}_2) \\
\frac{\partial \mathbf{e}_2}{\partial \mathbf{X}} &= -\left[-\frac{\mathbf{S}(\mathbf{e}_1)}{\|\mathbf{e}_1 * \mathbf{r}\|} + 2\frac{(\mathbf{e}_1 * \mathbf{r})(\mathbf{S}(\mathbf{e}_1)(\mathbf{e}_1 * \mathbf{r}))^T}{\|\mathbf{e}_1 * \mathbf{r}\|^3} \;\; \mathbf{0}_{3\times 6}\right] \\
\frac{\mathrm{d}\mathbf{e}_1}{\mathrm{d}\mathbf{X}} &= \frac{\partial \mathbf{e}_1}{\partial \mathbf{X}_r}\frac{\partial \mathbf{X}_r}{\partial \mathbf{X}} \\
\frac{\mathrm{d}\mathbf{e}_2}{\mathrm{d}\mathbf{X}} &= \frac{\partial \mathbf{e}_2}{\partial \mathbf{X}} + \frac{\partial \mathbf{e}_2}{\partial \mathbf{e}_1}\frac{\mathrm{d}\mathbf{e}_1}{\mathrm{d}\mathbf{X}} \\
\frac{\mathrm{d}\mathbf{e}_3}{\mathrm{d}\mathbf{X}} &= \frac{\partial \mathbf{e}_3}{\partial \mathbf{e}_1}\frac{\mathrm{d}\mathbf{e}_1}{\mathrm{d}\mathbf{X}}.
\end{aligned}
\tag{A.4}
$$

Partials are then taken for each of the terms in the equation of motion A.2

$$
\begin{aligned}
\mathbf{T}_a &= \boldsymbol{\omega} * \mathbf{a} \\
\frac{\partial \mathbf{T}_a}{\partial \mathbf{X}} &= [\mathbf{0}_{3\times 6} \;\; \mathbf{S}(\boldsymbol{\omega})] \\
\frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}\mathbf{X}_r} &= \frac{1}{\|\dot{\mathbf{r}}_r\|^2}\left[-\mathbf{S}(\ddot{\mathbf{r}}_r) \;\; \mathbf{S}(\dot{\mathbf{r}}_r)\right] - 2\frac{\dot{\mathbf{r}}_r * \ddot{\mathbf{r}}_r}{\|\dot{\mathbf{r}}_r\|^4}\left[\dot{\mathbf{r}}_r^T \;\; \mathbf{0}_{1\times 3}\right] \\
\frac{\partial \mathbf{T}_a}{\partial \boldsymbol{\omega}} &= -\mathbf{S}(\mathbf{a}) \\
\frac{\mathrm{d}\mathbf{T}_a}{\mathrm{d}\mathbf{X}} &= \frac{\partial \mathbf{T}_a}{\partial \mathbf{X}} + \frac{\partial \mathbf{T}_a}{\partial \boldsymbol{\omega}}\frac{\mathrm{d}\boldsymbol{\omega}}{\mathrm{d}\mathbf{X}_r}\frac{\mathrm{d}\mathbf{X}_r}{\mathrm{d}\mathbf{X}},
\end{aligned}
\tag{A.5}
$$

(A.6)

$$
\begin{aligned}
\mathbf{T}_b &= \dot{\varphi}\mathbf{e}_1 * \mathbf{a} \\
\frac{\partial \mathbf{T}_b}{\partial \mathbf{X}} &= \dot{\varphi}\left[\mathbf{0}_{3\times 6}\ \mathbf{S}(\mathbf{e}_1)\right] \\
\frac{\partial \mathbf{T}_b}{\partial \mathbf{e}_1} &= -\mathbf{S}(\mathbf{e}_1) \\
\frac{\mathrm{d}\mathbf{T}_b}{\mathrm{d}\mathbf{X}} &= \frac{\partial \mathbf{T}_b}{\partial \mathbf{X}} + \frac{\partial \mathbf{T}_b}{\partial \mathbf{e}_1}\frac{\mathrm{d}\mathbf{e}_1}{\mathrm{d}\mathbf{X}_r}\frac{\mathrm{d}\mathbf{X}_r}{\mathrm{d}\mathbf{X}},
\end{aligned}
\tag{A.7}
$$

$$
\mathbf{T}_c = -\dot{D}\mathbf{e}_1 \tag{A.8}
$$

$$
\frac{\mathrm{d}\mathbf{T}_c}{\mathrm{d}\mathbf{X}} = -\mathbf{e}_1\frac{\mathrm{d}\dot{D}}{\mathrm{d}\mathbf{X}} - \dot{D}\frac{\mathrm{d}\mathbf{e}_1}{\mathrm{d}\mathbf{X}}, \tag{A.9}
$$

$$
\begin{aligned}
\mathbf{T}_d &= \dot{L}(-\mathbf{e}_2\sin(\varphi) + \mathbf{e}_3\cos(\varphi)) \\
\frac{\mathrm{d}\mathbf{T}_d}{\mathrm{d}\mathbf{X}} &= (-\mathbf{e}_2\sin(\varphi) + \mathbf{e}_3\cos(\varphi))\frac{\mathrm{d}\dot{L}}{\mathrm{d}\mathbf{X}} + \\
&\quad \dot{L}(-\mathbf{e}_2\sin(\varphi) + \mathbf{e}_3\cos(\varphi))
\end{aligned}
\tag{A.10}
$$

,

where appropriate expressions of $\dot{L}$ and $\dot{D}$ are given by

$$
\dot{L} = \frac{S}{m}\left(\dot{C}_L q + C_L\dot{q}\right) \tag{A.11}
$$

$$
\dot{D} = \frac{S}{m}\left(\dot{C}_D q + C_D\dot{q}\right) \tag{A.12}
$$

with

$$
\begin{aligned}
q &= \frac{1}{2}\rho\|\dot{\mathbf{r}}_r\|^2 && \text{(A.13)}\\
\dot{q} &= \frac{1}{2}\dot{\rho}\|\dot{\mathbf{r}}_r\|^2 + \rho\ddot{\mathbf{r}}_r * \dot{\mathbf{r}}_r, && \text{(A.14)}
\end{aligned}
$$

Hence, we have the following partials:

$$
\begin{aligned}
\frac{\partial q}{\partial \mathbf{X}} &= \frac{1}{2}\frac{\mathrm{d}\rho}{\mathrm{d}\mathbf{X}}\|\dot{\mathbf{r}}_r\|^2\\
\frac{\mathrm{d}q}{\mathrm{d}\mathbf{X}_r} &= \rho\left[\dot{\mathbf{r}}_r^T\ \mathbf{0}_{1\times 3}\right] && \text{(A.15)}\\
\frac{\mathrm{d}q}{\mathrm{d}\mathbf{X}} &= \frac{\partial q}{\partial \mathbf{X}} + \frac{\mathrm{d}q}{\mathrm{d}\mathbf{X}_r}\frac{\mathrm{d}\mathbf{X}_r}{\mathrm{d}\mathbf{X}}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \dot{q}}{\partial \mathbf{X}} &= \frac{1}{2}\frac{\mathrm{d}\dot{\rho}}{\mathrm{d}\mathbf{X}}\|\dot{\mathbf{r}}_r\|^2 + \frac{\mathrm{d}\rho}{\mathrm{d}\mathbf{X}}\ddot{\mathbf{r}}_r \odot \dot{\mathbf{r}}_r\\
\frac{\mathrm{d}\dot{q}}{\mathrm{d}\mathbf{X}_r} &= \dot{\rho}\left[\dot{\mathbf{r}}_r\ \mathbf{0}_{1\times 3}\right] + \rho\left[\ddot{\mathbf{r}}_r^T\ \dot{\mathbf{r}}_r^T\right] && \text{(A.16)}\\
\frac{\mathrm{d}\dot{q}}{\mathrm{d}\mathbf{X}} &= \frac{\partial \dot{q}}{\partial \mathbf{X}} + \frac{\mathrm{d}\dot{q}}{\mathrm{d}\mathbf{X}_r}\frac{\mathrm{d}\mathbf{X}_r}{\mathrm{d}\mathbf{X}}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\mathrm{d}\dot{L}}{\mathrm{d}\mathbf{X}} &= \frac{S}{m}\left(C_L\frac{\mathrm{d}\dot{q}}{\mathrm{d}\mathbf{X}} + \dot{C}_L\frac{\mathrm{d}q}{\mathrm{d}\mathbf{X}}\right) && \text{(A.17)}\\
\frac{\mathrm{d}\dot{D}}{\mathrm{d}\mathbf{X}} &= \frac{S}{m}\left(C_D\frac{\mathrm{d}\dot{q}}{\mathrm{d}\mathbf{X}} + \dot{C}_D\frac{\mathrm{d}q}{\mathrm{d}\mathbf{X}}\right).
\end{aligned}
$$

To find the partial with respect to the parameters λ and φ we need a different representation for $\dot{L}$ and $\dot{D}$

$$\dot{D} = \beta_m q(1+\lambda^2)\left(\frac{2\ddot{\mathbf{r}}\odot\dot{\mathbf{r}}}{\dot{r}^2} - \frac{2(\boldsymbol{\Omega}*\mathbf{r})\odot\ddot{\mathbf{r}}_r}{\dot{r}_r^2} - \frac{\dot{\mathbf{r}}\odot\mathbf{r}}{H_o r}\right),$$

and

$$\dot{L} = 2\beta_m\left(\frac{C_L}{C_D}\right)_{\max} q\lambda\left(\frac{2\ddot{\mathbf{r}}\odot\dot{\mathbf{r}}}{\dot{r}_r^2} - \frac{2(\boldsymbol{\Omega}*\mathbf{r})\odot\ddot{\mathbf{r}}_r}{\dot{r}_r^2} - \frac{\dot{\mathbf{r}}\odot\mathbf{r}}{H_o r}\right).$$

Then we have

$$\frac{\partial\mathbf{T}_c}{\partial\lambda} = 2\beta_m q\left(\frac{2\ddot{\mathbf{r}}\odot\dot{\mathbf{r}}}{\dot{r}^2} - \frac{2(\boldsymbol{\Omega}*\mathbf{r})\odot\ddot{\mathbf{r}}_r}{\dot{r}_r^2} - \frac{\dot{\mathbf{r}}\odot\mathbf{r}}{H_o r}\right)\mathbf{e}_1 \tag{A.18}$$

$$\frac{\partial\mathbf{T}_d}{\partial\lambda} = 4\beta_m q\left(\frac{2\ddot{\mathbf{r}}\odot\dot{\mathbf{r}}}{\dot{r}^2} - \frac{2(\boldsymbol{\Omega}*\mathbf{r})\odot\ddot{\mathbf{r}}_r}{\dot{r}_r^2} - \frac{\dot{\mathbf{r}}\odot\mathbf{r}}{H_o r}\right)(-\mathbf{e}_2\sin(\varphi)+\mathbf{e}_3\cos(\varphi)) \tag{A.19}$$

$$\frac{\partial\mathbf{T}_d}{\partial\varphi} = -\dot{L}(\mathbf{e}_2\cos(\varphi)+\mathbf{e}_3\sin(\varphi)) \tag{A.20}$$

Finally, the state propagation partial matrix $\mathbf{F}$ is given by

$$\mathbf{F} = \left[\begin{array}{ccc:c:c} \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} \\ & \frac{\mathrm{d}\mathbf{T}_a}{\mathrm{d}\mathbf{X}}+\frac{\mathrm{d}\mathbf{T}_b}{\mathrm{d}\mathbf{X}}+\frac{\mathrm{d}\mathbf{T}_c}{\mathrm{d}\mathbf{X}}+\frac{\mathrm{d}\mathbf{T}_d}{\mathrm{d}\mathbf{X}} & & \frac{\partial\mathbf{T}_c}{\partial\lambda}+\frac{\partial\mathbf{T}_d}{\partial\lambda} & \frac{\partial\mathbf{T}_d}{\partial\varphi} \\ \mathbf{0}_{2\times3} & \mathbf{0}_{2\times3} & \mathbf{0}_{2\times3} & \mathbf{0}_{2\times1} & \mathbf{0}_{2\times1} \end{array}\right]. \tag{A.21}$$

REFERENCES

[1] Tobak M. Peterson, V. L., "Angle-of-attack convergence of spinning bodies entering planetary atmospheres at large inclinations to the flight path", *Ames Research Center,NASA-TR-R-210*, p. 21, 1964.

[2] R.L. Nelson, "Angle-of-attack convergence of spinning bodies entering planetary atmospheres at large inclinations to the flight path", *Langley Aeronautical Laboratory,NACA TN 3737*, p. 51, 1964.

[3] W.R. Chadwick and P. Zarchan, "Interception of spiraling ballistic missiles", in *American Control Conference, 1995. Proceedings of the*, jun 1995, vol. 6, pp. 4476 –4483 vol.6.

[4] P. Zarchan, "Tracking and intercepting spiraling ballistic missiles", in *Position Location and Navigation Symposium, IEEE 2000*, 2000, pp. 277 –284.

[5] O. Dubois-Matra, *Development of Multisensor Fusion Techniques with Gating Networks Applied to Reentry Vehicles*, PhD thesis, University of Texas, Austin, may 2003.

[6] R. H. Bishop and A. C. Antoulas, "A nonlinear approach to the aircraft tracking problem", *AIAA Journal of Guidance, Control, and Dynamics*, vol. 17, no. 5, pp. 1124–1130, 1994.

[7] Kim Jinwhan., Menon P. K., and E Ohlmeyer, "Motion models for use with the maneuvering ballistic missile tracking estimators", in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, August 2010.

[8] Kim Jinwhan., Menon P. K., and E Ohlmeyer, "Tracking of spiraling reentry vehicles with varying frequency using the unscented kalman filter", in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, August 2010.

[9] Arnaud Doucet, Nando de Freitas, and Neil Gordon, *Sequential Monte Carlo Methods In Practice*, Springer, New York, 2001.

[10] J.D. Murray, *Mathematical Biology: I. An Introduction (Interdisciplinary Applied Mathematics)(Pt. 1)*, New York: Springer, 2007.

[11] karl Astrom, *Introdution To Stochastic Control Theory*, Dover Publications, Mineola, N.Y, 2006.

[12] Sebastian Thrun, *Probabilistic robotics*, MIT Press, Cambridge, Mass, 2005.

[13] J. Regan, Frank and M. Anandakrishnan, Satya, *Dynamics of Atmospheric Re-Entry*, AIAA Education Series, American Institute of Aeronautics and Astronautics, 1993.

[14] B. O'Neill, *Elementary Differential Geometry*, Academic Press, INC., Orlando, Florida, 1966.

[15] Weng Wai Leong, "Trade-off study for the hit-to-kill interception of ballistic missiles in the boost phase", Master's thesis, Naval Postgraduate School, Monterey, California, 2009.

[16] Y. Ho and R. Lee, "A bayesian approach to problems in stochastic estimation and control", *Automatic Control, IEEE Transactions on*, vol. 9, no. 4, pp. 333 – 339, oct 1964.

[17] Andrew H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, apr 1970.

[18] R L Stratonovich, *Conditional Markov Processes And Their Application To The Theory Of Optimal Control*, American Elsevier Publ CO, 1968.

[19] Rudolph Emil Kalman, "A new approach to linear filtering and prediction problems", *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[20] H. W. Sorenson and D. L. Alspach, "Recursive bayesian estimation using gaussian sums", *Automatica*, vol. 7, no. 4, pp. 465–479, jul 1971.

[21] Genshiro Kitagawa, "Non-Gaussian State-Space Modeling of Nonstationary Time Series", *Journal of the American Statistical Association*, vol. 82, no. 400, pp. 1032–1041, 1987.

[22] J. Handschin, "Monte Carlo Techniques for Prediction and Filtering of Non-Linear Stochastic Processes", *Automatica*, vol. 6, no. 4, pp. 555–563, jul 1970.

[23] D. Q. Mayne and J. E. Handschin, "Monte carlo techniques to estimate the conditional expectation in multi-stage non-linear filtering", *International Journal of Control*, vol. 5, no. 5, pp. 547–559, 1969.

[24] N.J. Gordon, D.J. Salmond, and A.F.M. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation", *Radar and Signal Processing, IEEE Proceedings*, vol. 140, no. 2, pp. 107–113, apr 1993.

[25] Athanasios Papoulis, *Probability, random variables, and stochastic processes*, McGraw-Hill, Boston, 2002.

[26] Jesper Carlsson, Kyoung-Sook Moon, Anders Szepessy, and Ral Tempone Georgios Zouraris, "Stochastic differential equations: Models and numerics", 2010.

[27] E.A. Wan and R. van der Merwe, "The unscented kalman filter for nonlinear estimation", in *Proc. of the IEEE Symp. on Adaptive Systems for Signal Processing, Communication and Control (AS-SPCC)*, Lake Louise, Alberta, Canada, 2000.

[28] Dongming Zhao, Qingbin Wang, Huan Bao, and Zhan Gao, "The second order central divided-difference kalman filter.", 2011, pp. 2637–2640, IEEE.

[29] Subhash Challa, Yaakov Bar-Shalom, and Vikram Krishnamurthy, "Nonlinear filtering via generalized edgeworth series and gauss-hermite quadrature.", *IEEE Transactions on Signal Processing*, vol. 48, no. 6, pp. 1816–1820, 2000.

[30] Juha Sarmavuori and Simo Srkk, "Fourier-hermite kalman filter.", *IEEE Trans. Automat. Contr.*, vol. 57, no. 6, pp. 1511–1515, 2012.

[31] Rudolph van der Merwe, Nando de Freitas, Arnaud Doucet, and Eric Wan, "The Unscented Particle Filter", in *Advances in Neural Information Processing Systems 13*, nov 2001.

[32] Shaolin Lv, Jiabin Chen, and Zhide Liu, "Udut continuous-discrete unscented kalman filtering", in *Proceedings of the 2008 Second International Symposium on Intelligent*

Information Technology Application - Volume 02, Washington, DC, USA, 2008, IITA '08, pp. 876–879, IEEE Computer Society.

[33] Jianxin Wu, "Some properties of the gaussian distribution", 2004.

[34] Simo Srkk, "Recursive bayesian inference on stochastic differential equations", Tech. Rep., Helsinki University of Technology, 2006.

[35] William Brogan, *Modern control theory*, Prentice Hall, Englewood Cliffs, N.J, 1991.

[36] Samuel Blackman, *Design and analysis of modern tracking systems*, Artech House, Boston, 1999.

[37] Dirk Kroese, *Handbook of Monte Carlo methods*, Wiley, Hoboken, N.J, 2011.

[38] Anton Haug, *Bayesian Estimation and Tracking a Practical Guide*, John Wiley and Sons, Hoboken, 2012.

[39] Jeroen D. Hol, Thomas B. Schon, and Fredrik Gustafsson, "On resampling algorithms for particle filters", in *Nonlinear Statistical Signal Processing Workshop, 2006 IEEE*, sept. 2006, pp. 79 –82.

[40] C.S. Agate and K.J. Sullivan, "Particle filtering for tracking multiple, road-constrained targets", in *Proceedings of the SPIE Conference on Signal Processing, Sensor Fusion, and Target Recognition XII*, Apr. 2003, vol. 5096, pp. 256–266.

[41] R. H. Bishop, "Development of a spiraling marv threat model", Nov 1999.

[42] R. H. Bishop and O. Dubois-Matra, "Observability analysis and parameter estimation", June 5 2000.

[43] R.H. Bishop and A.C. Antoulas, "Nonlinear approach to aircraft tracking problem", *AIAA Journal of Guidance, Control, and Dynamics*, vol. 17, no. 5, pp. 1124–1130, 1994.

[44] I.S. Gradshteyn and I.M. Ryzhik., *Table of Integrals, Series and Products*, Academic Press, NY., 1965.

[45] Dubois-Matra O. Bishop, R. H. and T. Ely, "Robust entry navigation using hierarchical filter architectures regulated with gating networks", *16th International Symposium on Space Flight Dynamics*, vol. 17, no. 5, pp. 1124–1130, 2001.

[46] Olivier Dubois-Matra and R. H. Bishop, "Tracking and identification of a maneuvering reentry vehicle", *AIAA Journal of Guidance, Control, and Dynamics*, vol. 17, no. 5, pp. 1124–1130, 2003.

[47] Geoffrey Taylor, Lindsay Kleeman, Geoffrey Taylor, and Lindsay Kleeman, "C iterated extended kalman filter", in *Visual Perception and Robotic Manipulation*, vol. 26 of *Springer Tracts in Advanced Robotics*, pp. 185–187. Springer Berlin / Heidelberg, 2006.

[48] Spencer David A., Blanchard Robert C., Thurman Sam W., Braun Robert D., Peng Chia-Yen, and Kallemeyn Pieter H., "Mars pathfinder atmospheric entry reconstruction", Tech. Rep., 1998.

[49] R.D. Braun and R.M. Manning, "Mars exploration entry, descent and landing challenges", in *Aerospace Conference, 2006 IEEE*. IEEE.