RADAR TARGET-TRACKING AND MEASUREMENT-ORIGIN UNCERTAINTY

RADAR TARGET-TRACKING AND MEASUREMENT-ORIGIN UNCERTAINTY

BY

EDUARDO SANTOS DIAZ. B.ENG., M. SC.

A THESIS
SUBMITTED TO THE SCHOOL OF GRADUATE STUDIES
OF MCMASTER UNIVERSITY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

McMaster University DOCTOR OF PHILOSOPHY (2018) Hamilton, Ontario
(School of Computational Science and Engineering)
TITLE: Radar Target-tracking and Measurement-origin Uncertainty
AUTHOR: Eduardo Santos Diaz, B.Eng. M. Sc.
SUPERVISOR: Professor Simon Haykin
NUMBER OF PAGES: ix, 110

To my parents, Eduardo and Maria Eugenia, who through their sacrifice made me who I am today, for their unconditional love, and a life filled with happiness.

To my brother, Alejandro, for always being there for me and making this part of my journey so much easier.

Abstract

Target tracking refers to the process of estimating the state of a moving object from remote and noisy measurements. In this thesis we consider the Bayesian filtering framework to perform target tracking under nonlinear models, a target moving in continuous time, and measurements that are available in discrete time intervals (known as continuous-discrete). The Bayesian filtering theory establishes the mathematical basis to obtain the posterior probability density function of the state, given the measurement history. This probability density function contains all the information required about the state of the target. It is well documented that there is no exact solution for posterior density under the models mentioned. Hence, the approximation of such density functions have been studied for over four decades. The literature demonstrates that this has led to the development of multiple filters.

In target tracking, due to the remote sensing performed, an additional complication emerges. The measurements received are not always from the desired target and could have originated from unknown sources, thus making the tracking more difficult. This problem is known as a measurement origin uncertainty. Additionally to the filters, different methods have been proposed to address the measurement origin uncertainty due to its negative impact, which could cause a false track. Unfortunately, a final solution has yet to be achieved.

The first proposal of this thesis is a new approximate Bayesian filter for continuous-discrete systems. The new filter is a higher accuracy version of the cubature Kalman filter. This filter is developed using a fifth-degree spherical-radial cubature rule and the Ito-Taylor expansion of order 1.5 for dealing with

stochastic differential equations. The second proposal is an improved version of the probabilistic data association method. The proposed method utilizes the maximum likelihood values for selecting the measurements that are used for the data association. In the first experiment, the new filter is tested in a challenging 3-dimensional turn model, demonstrating superiority over other existing filters. In a second and third experiments, the proposed data association method is tested for target tracking in a 2-dimensional scenarios under heavy measurement origin uncertainty conditions. The second and third experiments demonstrate the superiority of the proposed data association method compared to the probabilistic data association.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Simon Haykin, for giving me the opportunity to pursue my Ph.D. degree under his guidance. During these four years, he has been the utmost help and inspiration and unceasingly encouraged me to move forward. It has been a privilege to work by his side and gain such extensive knowledge from him.

I would also like to thank the two members of the committee - Dr. Paul Faure and Dr. Bartosz Protas for providing me with their valuable support and helpful feedback in this important stage of my academic career.

Special thanks to Dr. Thomas Hurd for collaborating with me during the last part of this project. It was both an honor and a delight to work with him.

Additionally, I would like to thank Dr. Hugh Griffiths for the assessment of my dissertation and for his very kind words about my work.

It is also very important for me to thank Dr. Aldo Orozco Lugo from CINVESTAV (Mexico) for giving me this tremendous opportunity by helping me to come to Canada. Without his help, none of this would be possible.

Many thanks to all my beloved friends for the time spent. These four amazing years will remain in my heart for many years to come.

Last but not least, I would like to thank CONACYT (Mexico) for providing me with the funding to pursue my degree through a scholarship granted (CVU: 231857).

Contents

List of Figures

List of Abbreviations

3D-IT-CKF Third-Degree Ito-Taylor Cubature Kalman Filter

3D-RK-CKF Third-Degree Runga-Kutta Cubature Kalman Filter

5D-CKF Fifth-Degree Cubature Kalman Filter

5D-IT-CKF Fifth-Degree Ito-Taylor Cubature Kalman Filter

5D-RK-CKF Fifth-Degree Runga-Kutta Cubature Kalman Filter

CD-5D-CKF Continuous-Discrete Fifth-Degree Cubature Kalman Filter

CD-CKF Continuous-Discrete Cubature Kalman Filter

CD-SCKF Continuous-Discrete Square-root Cubature Kalman Filter

CDS Cognitive Dynamic Systems

CKF Cubature Kalman Filter

EKF Extended Kalman Filter

JPDA Joint Probability Data Association

MHT Multiple Hypothesis Tracker

ML Maximum Likelihood

MLDA Maximum Likelihood Data Association

MMSE Minimum Mean Square Error

NASA National Aeronautics and Space Administration

NNF Nearesth-Neighbor Filter

PDA Probabilistic Data Association

PF Particle Filter

RMSE Root-Mean Square Error

SDE Stochastic Differential Equation

SNF Strongest-Neighbor Filter

Chapter 1

Introduction

1.1 Problem Statement

The term filtering conventionally refers to the action of extracting information about the state of a time-varying system from noisy data to "filter out" the noise (Bar-Shalom et al., 2004). In order to filter the noise, an "optimal" statistical error criterion are established. Optimal filtering is then the class of methods that can be used for estimating the state of a system based on these statistical error criteria.

Filtering has been studied since the early work of Kolmogorov (1939) in the beginning of the 20th century . Due to the presence of noise in nature, a wide spectrum of applications made use of this theory, including communications, radar, navigation, biomedical engineering, weather prediction and financial engineering among others (Haykin, 2008; Särkkä, 2013). In this thesis I focus on radar applications, where the state of a system is usually described by the collection of dynamic variables such as position, velocity, angular velocity and

range.

For this purpose, I have adopted the framework of Bayesian filtering, which refers to the Bayesian way of formulating optimal filtering. The roots of Bayesian filtering were developed in optimal linear filtering. The first optimal recursive estimators were developed for linear systems due to their mathematically simplicity, and the optimal criterion was the least square optimality. The Wiener filter is one of the pioneering methods in which least squares filtering is applied to Gaussian signals (Wiener, 1949).

The growth of interest by the scientific community for filtering happened in a seminal article with the discovery of the Kalman filter (Kalman et al., 1960). The Kalman filter provides the solution to the minimum mean square error (MMSE) for linear discrete-time systems, which will be described in section 2.2 of this thesis. The relevance of the Kalman filter was recognized by the National Aeronautics and Space Administration (NASA) and, due to its applicability, was implemented in the navigation computer of the Apollo mission (Grewal and Andrews, 2010).

Later on, the Kalman filter was discovered to be a special case of a broader class of filters known as Bayesian filters (Ho and Lee, 1964). Indeed, the Bayesian filter established a general unifying framework for sequential state estimation. Despite the fact that the original Kalman filter was based on the least squares approach, the same equations can be obtained from a purely probabilistic Bayesian analysis (Särkkä, 2013).

In radar applications, recursive filters like the Kalman filter are used to estimate the state of a target from remote measurements; this has been defined as a target tracking. Target tracking faces many challenges, including nonlinearities in the stochastic dynamic system, noise, etc. For these reasons, nonlinear Kalman filters have been an area of interest for the radar community with no final solution achieved to date. In this thesis I focus on accurate target tracking, which incorporates the challenge of nonlinear filtering. That said, additional complications are presented and will be discussed in a later chapter.

To make things clear, Figure 1.1 presents a basic block diagram that illustrates the challenge of state estimation. As shown, the state estimator only has access to the system measurements, which are corrupted by measurement noise.

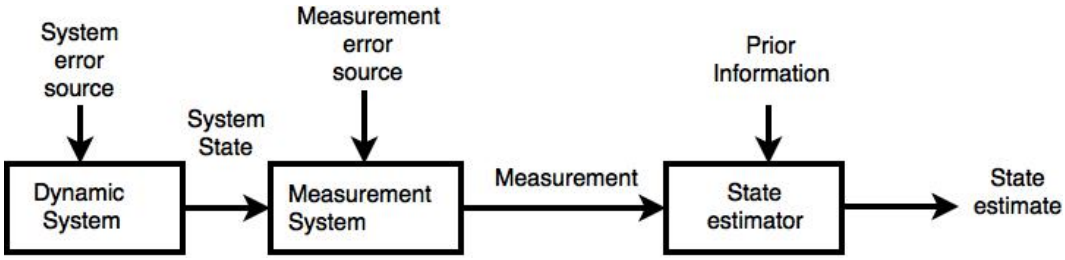

Figure 1.1: Simple block diagram of state estimation

The goal for the state estimator is to use knowledge about the dynamic system, via the measurement model, the probabilistic model of disturbances and the prior information, with the aim of getting the most accurate state estimate.

In radar, system measurements are usually echoes of a transmitted signal; therefore, the state estimation is presented with additional challenges, such as, selecting the transmission power, the type of wave form transmitted, the frequency of the signal transmitted, presence of clutter, etc.

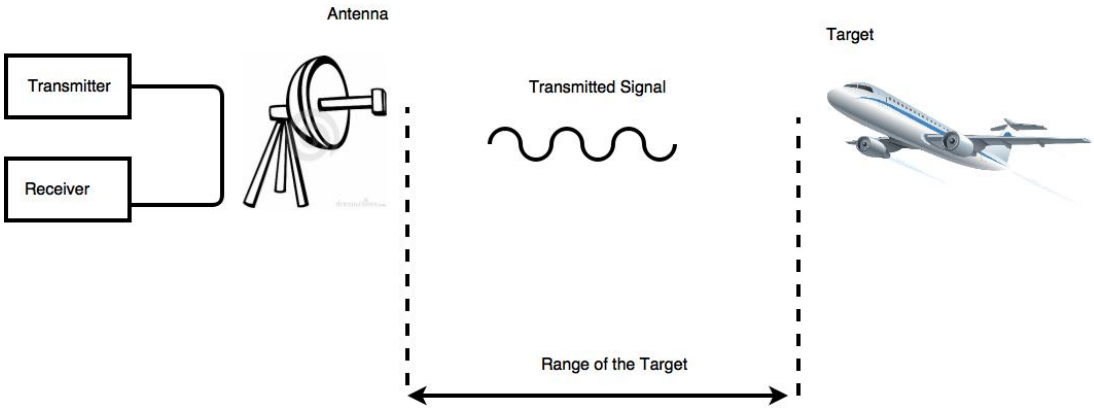

Figure 1.2: Basic block diagram of radar operation. The range of the target can be obtained through the time it takes the signal to travel forward to the target and back to the antenna.

The basic operation principle of radar is depicted in Figure 1.2, where the range (R) to the target can be obtained with the equation

$$R = \frac{cT_R}{2}, \tag{1.1}$$

where c is the transmission speed of signal in medium and T_R is the propagation time required for the radar signal to travel to the target and back to the antenna.

Because of the complexity of the system, radar target tracking has been divided in two main blocks: target detection and target tracking.

Target detection is the algorithm responsible of separating the echo of the transmitted signals originating from the environment and the echo of the transmitted signals originating from the target. In this manuscript, I focus mainly on target tracking and assume that target detection has been correctly performed by the system; however, I am aware that detection algorithms can not completely discriminate whether a signal originated from a target or the environment.

Therefore, we must assume that there are false alarms (false target detections) in the system. The presence of these false alarms generates a problem known as a measurement origin uncertainty, and it occurs when the tracking algorithm receives measurements that could have been generated by the environment or target surrounding. Unfortunately, filters are not designed for dealing with clutter, and algorithms have to be applied in conjunction with filters to achieve correct functioning.

The challenge ahead that remains, is to execute precise target tracking, focusing on two ideas that have a significant impact on the accuracy of target tracking. First, I address the approximate Bayesian filters by developing the most accurate and practical filter possible. Second, I further develop the techniques for dealing with measurement origin uncertainty.

1.2 Contributions

There are two main contribution of this thesis:
(1) The first contribution is the theoretical development of a new nonlinear Bayesian filter for systems with continuous-time dynamics and measurements in discrete time. The name of the filter is the continuous-discrete fifth-degree cubature Kalman filter (CD-5D-CKF), and it is an extension of the successful cubature Kalman filter first developed at McMaster (Arasaratnam and Haykin, 2009).

The new filter I developed is based on the theory used for the development of the cubature Kalman filter, in which a spherical-radial cubature rule is

combined with a Gauss-Laguerre quadrature rule. In addition, the filter uses two different approaches for solving stochastic differential equations presented in systems with continuous-time dynamics. The computational complexity is polynomial and therefore can be applied to high-dimensional nonlinear filtering problems.

(2) The second contribution is the proposed modified probabilistic data association, which is a technique for tracking a taget with clutter conditions. The new algorithm provides a substantial increase in tolerance against high clutter environments compared to the well established PDA.

The proposed technique does not requires any special adjustment for different clutter densities. Moreover the computational complexity and computational time is similar to other practical techniques as the PDA, which is an important concern for real systems.

Finally, I validate the formulation of the new filter and the new modified PDA through a number of challenging computer simulations.

1.3 Organization of Thesis

Chapter 2 reviews the literature on discrete-time nonlinear filters. Specifically, we focus on the different approaches for solving the nonlinear integrals presented in the Bayesian filter.

Chapter 3 reviews the theoretical approach for filtering that operates with

continuous-time dynamics and discrete measurements.

In Chapter 4, the new continuous-discrete fifth-degree cubature Kalman filter is developed using the theory presented in chapter 3 and a fifth-degree spherical-radial cubature rule. Additionally, computer experiments are shown to demonstrated the superiority of the proposed filters against the traditional continuous-discrete cubature Kalman filter.

In Chapter 5 we review the problem measurement origin uncertainty, the modelling of clutter in target tracking and detailed the popular probabilistic data association.

Chapter 6 introduces the development of the modified probabilistic data association. In addition, two computer experiments show the superiority of the proposed method compared to the probabilistic data association.

Chapter 7 gives a brief summary of the thesis and proposes future topics of research.

In Appendix A, the formulation of a two joint Gaussian random variables is depicted.

In Appendix B, the conditional distribution of a joint Gaussian random variable is derived.

Chapter 2

Literature Review

Discrete-time filtering has been widely studied by the scientific community compared to continuous-time filtering (Crouse, 2015). Discrete-time filters have been applied to most real-life applications; even most of the dynamics and assumptions of real-life systems are in continuous-time. The are two main justifications for the use of discrete-time filters. First, the individual measurements from a real-life systems are collected in discrete time, so in that sense, the discrete-time assumption is correct. Second, the mathematics is simpler because diference equations are used instead of differential equations which are more complex. By virtue of the relevant work in discrete-time filtering theory, filters developed for dynamics in continuous-time are usually additions or expansions of filters conceived for discrete-time. For that reason, this chapter focuses on discrete-time filtering theory to provide the bases for the explanation of filters with continuous-time dynamics later in this thesis.

The discrete-time filtering theory considers the problem of sequentially estimat-

ing the state of the system $x_k \in \mathbb{R}^{nx}$ which evolves in time according to the equation

$$x_k = f(x_{k-1}) + v_{k-1}, \tag{2.1}$$

where the term $\mathbb{R}$ refers to real numbers and nx is the dimension of x_k. The state is observed through the noisy measurement via

$$z_k = h(x_k) + w_k, \tag{2.2}$$

where $z_k \in \mathbb{R}^{nz}$ is the measurement with dimension nz that is independent of nx; $f : \mathbb{R}^{nx} \to \mathbb{R}^{nx}$ is a nonlinear function and $h : \mathbb{R}^{nx} \to \mathbb{R}^{nz}$ is the measurement function; v_{k-1} and w_k are the uncorrelated process and measurement Gaussian noise sequences with zero means and covariances Q_{k-1} and R_k, respectively.

As mentioned in the previous chapter, the Bayesian filter provides a mathematical framework for sequential state estimation. Before giving further details about this filter, there is the need to explain its methodology, the assumptions endorsed, and why they are needed. The purpose of filtering is to calculate the value of the state of a system $(x_0, x_1, ..., x_k)$ at certain times $(t_0, t_1, ..., t_k)$ observed through measurements $(z_1, z_2, ..., z_k)$. This task is considered to be a statistical inversion problem. The solution is given by the joint posterior distribution of all the states given all the measurements, which can be calculated using Bayes' rule

$$p(x_{0:k}|z_{1:k}) = \frac{p(z_{1:k}|x_{0:k})p(x_{0:k})}{p(z_{1:k})}, \tag{2.3}$$

where $x_{0:k}$ is the state of the system from time 0 to time k; $p(x_{0:k})$ is the

prior probability distribution defined by the dynamic model, $p(z_{1:k}|x_{0:k})$ is the likelihood model, that is, a function used in statistics to describe the plausibility of the parameters in $x_{0:k}$, and $p(z_{1:k})$ is the normalization constant defined as

$$p(z_{1:k}) = \int p(z_{1:k})p(x_{0:k})dx_{0:k}\,. \tag{2.4}$$

This formulation, although completely correct, suffers the main disadvantage that every time a new measurement is received the full posterior distribution has to be reestimated. In dynamic systems; where the discrete measurements are received every sampling time, the dimensionality of the full posterior distribution increases with every measurement received, due to the fact that the joint posterior distribution of all the states is computed from time zero to the present state, leading to an intractable computation that is not feasible for applications. Knowing this and with the objective of finding a solution, instead of computing the full posterior distribution, we can calculate the marginal distribution. A marginal distribution is defined as the probability distribution of a single variable, or combination of variables in a multivariate distribution. Obtained from the multivariate distribution by integrating over the other variables (Everitt and Skrondal, 2002). In this case, the marginal distribution objective $p(x_k|z_{1:k})$ is the probability density function of the current state x_k given the set of previous and current measurements $z_{1:k} = (z_1, z_2, ..., z_k)$. The requirement for this approach is to limit the application to a Markov process (Dynkin, 1965). The Bayesian filter is the procedure to compute this marginal distribution and it is explained next.

2.1 Bayesian Filter

The five required elements for the description of the Bayesian filter are:

1. The initial distribution $p(x_0)$ that defines the initial knowledge of the state x_0 at the initial time step $k = 0$.

2. The dynamic model that describes the system dynamics and uncertainties as a Markov sequence defined as a transition probability distribution $p(x_k|x_{k-1})$.

3. The measurement model that is responsible for relating the current state x_k given the measurement z_k. This model is used to obtain the conditional probability distribution of the state given the measurement $p(z_k|x_k)$.

4. The joint probability distribution of the current and previous state $p(x_k, x_{k-1}|z_{1:k-1})$ given the previous set of measurements $z_{1:k-1} = \{z_1, ..., z_k - 1\}$.

5. The filter distribution $p(x_k|z_{1:k})$, which is computed by the Bayesian filter as the marginal distribution of the current state x_k given the current and previous set of measurements $z_{1:k} = \{z_1, ..., z_k\}$.

The computation of the Bayesian filter is performed in two basic steps, namely the time update and the measurement update, and both are explained next.

2.1.1 Time Update

The time update is obtained using Kolmogorov's forward equation (Jazwinski, 1970), where the old posterior probability density function is updated before receiving a new measurement. As a result, the predictive probability density function is obtained as it is shown by

$$\begin{aligned} p(x_k|z_{1:k-1}) &= \int_{\mathbb{R}^{nx}} p(x_k, x_{k-1}|z_{1:k-1})dx_{k-1} \\ &= \int_{\mathbb{R}^{nx}} p(x_{k-1}|z_{1:k-1})p(x_k|x_{k-1})dx_{k-1}, \end{aligned} \tag{2.5}$$

where $p(x_{k-1}|z_{1:k-1})$ is the old posterior distribution at time $(k-1)$, and the distribution $p(x_k|x_{k-1})$ is obtained from the dynamic model.

2.1.2 Measurement Update

The measurement update exploits the information about the current state contained in the newly received measurement. Using Bayes rule, the predictive density is updated and the posterior density of the current state is obtained as:

$$p(x_k|z_{1:k}) = \frac{1}{c_k} p(x_k|z_{1:k-1})p(z_k|x_k), \tag{2.6}$$

where $p(x_k|z_{1:k-1})$ was previously obtained in the time update, $p(z_k|x_k)$ can be obtained from the measurement equation, and the normalizing constant c_k is given by

$$\begin{aligned} c_k &= p(z_k|z_{1:k-1}) \\ &= \int_{\mathbb{R}^{nx}} p(x_k|z_{1:k-1})p(z_k|x_k)dx_k. \end{aligned} \tag{2.7}$$

The time update and measurement update operate in a cyclic basis, and a block diagram of the Bayesian filter is shown below in Figure 2.1

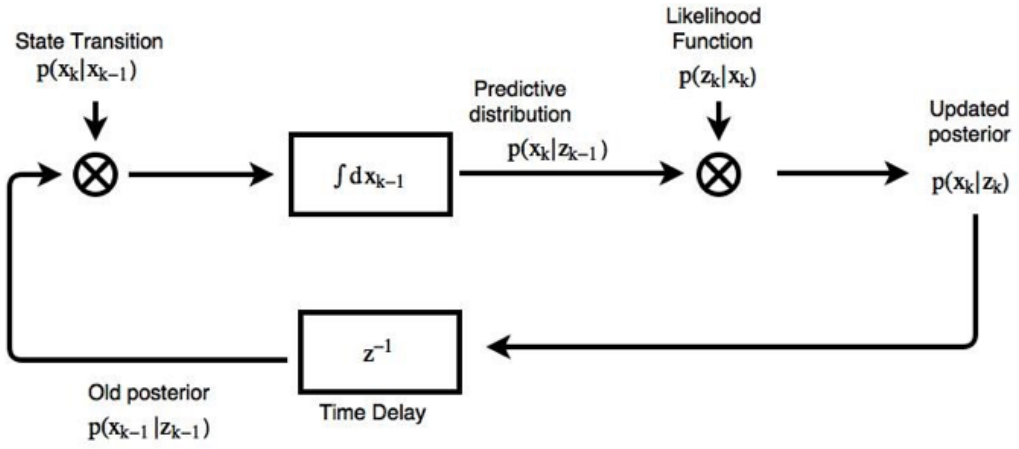

Figure 2.1: The Bayesian filter operates in a cyclic basis. The computed updated posterior becomes the old posterior after a time delay and the operations shown in equations (2.5),(2.6) and (2.7) have to be recomputed as shown in the cycle.

Depending on the assumptions of the system, the Bayesian filter may or may not have a closed form solution. The Kalman filter is a closed form solution, and it is due to the Gaussian and linear assumptions in the model. If the system is nonlinear, the integrals in the Bayesian filter are intractable and consequently, a numerical approximation must be used. Here are some examples of these numerical approximations:

- The extended Kalman filter (Jazwinski, 1970) uses a Taylor series expansion to linearize and approximate the nonlinearities in the dynamic models. The resulting estimated state is a Gaussian distribution.
- The unscented Kalman filter (Julier et al., 2000) computes the time and measurement update using the unscented transform to solve the nonlinear integrals, resulting in a Gaussian estimated state.

- The cubature Kalman filter (Arasaratnam and Haykin, 2009) uses the cubature rule to approximate the nonlinear integrals presented in the time and measurement updates. The result is also a Gaussian distribution.

- The Particle filters (Arulampalam et al., 2002) use a weighted set of samples to represent the posterior distribution. In this case, there is no need for a Gaussian approximation.

Owing to the relevance of the Kalman filter, we consider important to put its mathematical derivation in this thesis. Additionally, a brief description of the extended Kalman filter, the cubature Kalman and particle filters are presented in this chapter.

2.2 Kalman Filter

For the derivation of the Kalman filter, consider the discrete-time linear dynamic system described by the equation

$$x_k = F_{k-1}x_{k-1} + v_{k-1}, \tag{2.8}$$

where $x_k \in \mathbb{R}^{nx}$ is the state vector, F_{k-1} is a known and possibly time-varying matrix, $v_{k-1} \in \mathbb{R}^{nx}$ is the sequence of independent zero-mean white Gaussian process noise samples with covariance Q_k. The measurement equation is

$$z_k = H_k x_k + w_k, \tag{2.9}$$

where $z_k \in \mathbb{R}^{nz}$ is the measurement vector, H_k is a known matrix, $w_k \in \mathbb{R}^{nz}$ is the sequence of zero-mean white Gaussian measurement noise with covariance

R_k. The system described by these equations constitutes the linear Gaussian assumption. The notation use to refer to the true state of the system at time k is x_k whereas $\hat{x}_k$ refers to the state estimated. For the operation of the estimation algorithm, it is assumed that an initial estimate $\hat{x}_0$ of the true state x_0 and the initial estimated covariance P_0 are available. In probabilistic terms, we can describe the assumptions of linearity and Gaussianity as follows:

- The initial probability density function of the state is a known Gaussian distribution with mean $\hat{x}_0$ and covariance P_0. As it is shown by

$$p(x_0|z_0) \backsim N(\hat{x}_0, P_0). \tag{2.10}$$

- Due to the linearity of the system, the predicted state is a Gaussian distribution with mean given by the predicted state and covariance given by the process noise covariance. This is described by the equation

$$p(x_k|x_{k-1}) \backsim N(\hat{x}_{k|k-1}, Q_{k-1}). \tag{2.11}$$

- Due to the linearity in the measurement equation, the probability density function of the current measurement given the current state is a Gaussian distribution with mean given by the predicted measurement and covariance given by the measurement covariance.

$$p(z_k|x_k) \backsim N(z_{k|k-1}, R_k). \tag{2.12}$$

2.2.1 Time Update

The time update is responsible for the prediction of the probability density function of the old posterior. Due to the Gaussian assumption described, we can completely characterize the predicted state with the predicted mean and predicted covariance of the state with the following equations.

Using probability theory, we can obtain the predicted state as

$$\hat{x}_{k|k-1} = E[x_k|z_{k-1}] = E[F_{k-1}x_{k-1} + v_{k-1}]. \tag{2.13}$$

and considering that the process noise has a zero mean, we obtain

$$\hat{x}_{k|k-1} = F_{k-1}\hat{x}_{k-1|k-1}. \tag{2.14}$$

The predicted covariance is defined as:

$$\begin{aligned} P_{k|k-1} &= E\Big[[x_k - \hat{x}_{k|k-1}][x_k - \hat{x}_{k|k-1}]^T|z_k\Big] \\ &= F_{k-1}E\Big[[x_{k-1} - \hat{x}_{k-1}][x_{k-1} - \hat{x}_{k-1}]^T\Big]F_{k-1}^T + E[v_{k-1}v_{k-1}^T] \\ &= F_{k-1}P_{k-1|k-1}F_{k-1}^T + Q_{k-1}, \end{aligned} \tag{2.15}$$

where x_{k-1} refers to the true state of the system at time $k-1$ and $\hat{x}_{k-1}$ refers to the estimated state of the system at time $k-1$. Once the predicted state and predicted covariance have been obtained, the predicted density function $p(x_k|z_{1:k-1})$ can be characterized. Now we move into the measurement update.

2.2.2 Measurement Update

The predicted measurement $z_{k|k-1}$ that is required for the probability density function (2.12) is calculated as:

$$\begin{aligned} z_{k|k-1} &= E[H_k x_k + w_k | z_{k-1}] \\ &= HE[x_k + w_k | z_{k-1}]. \end{aligned} \tag{2.16}$$

Since w_k has a zero mean, the result is

$$z_{k|k-1} = H\hat{x}_{k|k-1}. \tag{2.17}$$

According to the measurement update in the Bayesian filter equation (2.6), a joint probability distribution is obtained by the multiplication of two Gaussian density functions as follows:

$$\begin{aligned} p(x_k, z_k | z_{1:k-1}) &= p(x_k | z_{1:k-1}) p(z_k | x_k) \\ &= N(z_k | H_k x_k, R_k) N(x_k | \hat{x}_{k|k-1}, P_{k|k-1}). \end{aligned} \tag{2.18}$$

As shown in Appendix A, the joint distribution obtained from the multiplication of two Gaussian distributions is

$$p(x_k, z_k) = N\left(\begin{bmatrix} x_k \\ z_k \end{bmatrix} \middle| \overline{x}, \overline{P} \right), \tag{2.19}$$

where

$$\overline{x} = \begin{bmatrix} \hat{x}_{k|k-1} \\ H_k \hat{x}_{k|k-1} \end{bmatrix},$$

and

$$\overline{P} = \begin{bmatrix} \hat{P}_{k|k-1} & \hat{P}_{k|k-1} H_k^T \\ H_k \hat{P}_{k|k-1} & H_k \hat{P}_{k|k-1} H^T + R_k \end{bmatrix}.$$

Once the joint probability distribution has been obtained, the conditional probability density function of x given z is given by the following equation,

$$p(x_k|z_{1:k}) = \frac{p(x_k, z_k|z_{1:k-1})}{p(z_k|z_{1:k-1})}. \tag{2.20}$$

Using the result in Appendix B, the conditional probability density function of x is given by

$$p(x_k|z_{1:k}) = N(x_k|\hat{x}_{k|k}, P_{k|k}), \tag{2.21}$$

where the mean and covariance are given by the equations

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} + [z_k - H_k \hat{x}_{k|k-1}], \tag{2.22}$$

$$P_{k|k} = P_{k|k-1} - P_{k|k-1} H_k^T (H_k P_{k|k-1} H^T + R_k)^{-}1 H_k P_{k|k-1}. \tag{2.23}$$

Altogether, these equations constitute the Kalman filter, which impacted sequential state estimation theory. Although its application is limited to the linear-Gaussian assumption, the Kalman filter boosted the development of a new set of filters based on its formulation. In many engineering applications the linearity assumption does not hold and the Kalman filter is not appropriate. Considering this, new approaches have been proposed to address the nonlinearities in a system. In this chapter, I present the three most relevant solutions for

dealing with system nonlinearities: extended Kalman filter (EKF), cubature Kalman filter (CKF) and Particle filter (PF). In the next sections, I describe each of these filtering approaches.

2.3 Extended Kalman Filter (EKF)

The EKF was proposed as one of the first solutions to operate for systems with nonlinear dynamics (Jazwinski, 1970). This filter is based on a linearization procedure using Taylor series expansion. For its description, we need to consider the nonlinear system described by equations (2.1) and (2.2). The linearization procedure consists of using the first or second order Taylor series of the nonlinear functions in the state-space model. Using the first order Taylor expansion we can obtain the following approximations:

$$f(x_{k-1}) \approx f(\hat{x}_{k-1|k-1}) + A_{k-1}\Big[x_{k-1} - \hat{x}_{k-1|k-1}\Big], \tag{2.24}$$

$$h(x_k) \approx h(\hat{x}_{k|k-1}) + B_k\Big[x_{k-1} - \hat{x}_{k|k-1}\Big], \tag{2.25}$$

where the values of A_{k-1} and B_k are matrices of partial derivatives known as Jacobians given by

$$A_{k-1} = \left.\frac{\partial f(x)}{\partial x}\right|_{x=\hat{x}_{k-1|k-1}}, \tag{2.26}$$

$$B_k = \left.\frac{\partial h(x)}{\partial x}\right|_{x=\hat{x}_{k|k-1}}. \tag{2.27}$$

Using the same procedure that was employed for the linear Kalman filter and the approximations (2.24) and (2.25), the time update and measurement update can computed as explained next.

2.3.1 Time Update

The predicted state is calculated using the expected value and the linearization approximation (2.24) over the old-posterior. This is shown by the equation

$$\begin{aligned}\hat{x}_{k|k-1} &= E[f(x_{k-1}) + v_{k-1}] \\ &\approx E\Big[f(\hat{x}_{k-1|k-1}) + A\Big[x_{k-1} - \hat{x}_{k-1|k-1}\Big] + v_k\Big] \\ &\approx f\Big(\hat{x}_{k-1|k-1}\Big).\end{aligned} \tag{2.28}$$

The predicted covariance is calculated with the same procedure used for the predicted covariance of the linear Kalman filter and is given by

$$\begin{aligned}P_{k|k-1} &= E[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T] \\ &\approx E\Bigg[\Big(f(\hat{x}_{k-1|k-1}) + A\Big[x_{k-1} - \hat{x}_{k-1|k-1}\Big] - f(\hat{x}_{k-1|k-1})\Big) \\ &\quad \Big(f(\hat{x}_{k-1|k-1}) + A\Big[x_{k-1} - \hat{x}_{k-1|k-1}\Big] - f(\hat{x}_{k-1|k-1})\Big)^T\Bigg] + E[v_{k-1}v_{k-1}^T] \\ &= A_{k-1}P_{k-1|k-1}A_{k-1}^T + Q_{k-1},\end{aligned} \tag{2.29}$$

where A_{k-1} is the Jacobian defined in equation (2.26). Once the predicted state and the predicted covariance have been obtained, we can move forward to compute the measurement update.

2.3.2 Measurement Update

The predicted measurement is calculated as:

$$\begin{aligned} z_{k|k-1} &\approx E\Big[h(x_k) + H_k\Big[x_k - \hat{x}_{k|k-1}\Big] + w_k\Big] \\ &= h(\hat{x}_{k|k-1}). \end{aligned} \tag{2.30}$$

The predicted measurement covariance is also known as the innovation covariance and is given by

$$\begin{aligned} S_k &= E\Big[\Big(h_k(x_k) - h(\hat{x}_{k|k-1})\Big)\Big(h_k(x_k) - h(\hat{x}_{k|k-1})\Big)^T\Big] \\ &= E\Big[\Big(h_k(\hat{x}_{k|k-1}) + B_k[x_{k-1} - \hat{x}_{k-1|k-1}] - h(\hat{x}_{k|k-1})\Big) \\ &\qquad \Big(h_k(\hat{x}_{k|k-1}) + B_k[x_{k-1} - \hat{x}_{k-1|k-1}] - h(\hat{x}_{k|k-1})\Big)^T + E[w_{k-1}w_{k-1}^T]\Big] \\ &= B_k P_{k|k-1} B_k^T + R_k. \end{aligned} \tag{2.31}$$

Once the predicted state, predicted measurement, predicted covariance, and innovation covariance have been calculated, using formulas from Appendix A the joint distribution of the predicted measurement and predicted state is given by:

$$p(x_k, z_k) = N\left(\begin{bmatrix} \hat{x}_{k|k-1} \\ \hat{z}_{k|k-1} \end{bmatrix}, \begin{bmatrix} P_{k|k-1} & P_{k|k-1}B^T \\ P_{k|k-1}^T B_k & B_k P_{k|k-1} B_k^T + R_k \end{bmatrix}\right). \tag{2.32}$$

With the joint distribution calculated, the updated measurement can be obtained with the marginal distribution; that is, the same equations used for the linear Kalman filter, but replacing the measurement matrix H_k with the

Jacobian B_k. The updated state of the EKF is then given by

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + P_{k|k-1}B_k^T(B_kP_{k|k-1}B_k^T + R_k)^{-1} + [z_k - h(\hat{x}_{k|k-1})], \quad (2.33)$$

and the corresponding covariance is given by

$$P_{k|k} = P_{k|k-1} - P_{k|k-1}B_k^T(H_kP_{k|k-1}B_k^T + R_k)^{-1}B_kP_{k|k-1}. \quad (2.34)$$

Althought the EKF has been one of the most relevant approaches for dealing with system nonlinearities and has been applied for many years, new theoretical advances have been made since its inception and improved calculation techniques are now available. For example, the CKF is one of the last and most relevant nonlinear filters, it has demonstrated better accuracy than other relevant nonlinear filters such as the EKF and the unscented Kalman filter (Sun and Tang, 2013). For this reason, the next section is devoted to the CKF.

2.4 Cubature Kalman Filter (CKF)

The CKF was developed as an approximate solution of the mathematical framework described by the Bayesian filter (Arasaratnam and Haykin, 2009). Observing that under the assumptions of Gaussianity and nonlinearity, the Bayesian filter reduces to computing multidimensional integrals of the form

$$\int (nonlinear function) \times (Gaussian function).$$

Arasaratnam and Haykin (2009) proposed the use of a third-degree spherical-radial cubature rule to approximate such an integral. Conductive to the use of the cubature rule, the integral composed by the nonlinear function and the Gaussian function had to be converted to spherical-radial coordinates and divided into two integrals. This is explained next.

Consider the integral composed by a nonlinear function $f(x)$ and Gaussian functions as

$$I(f) = \int_{\mathbb{R}^n} f(x) \exp(-x^T x) dx, \tag{2.35}$$

where the vector $x \in \mathbb{R}^n$. Utilizing the change of variables from the Cartesian coordinates to the radial vector r, let $x = rz$ with $z^T z = 1$ such that $x^T x = r^2$ for $r \in [0, \infty)$ and integral (2.35) becomes

$$I(f) = \int_0^\infty \int_{U_n} f(rz) r^{n-1} \exp(-r^2) d\sigma(z) dr, \tag{2.36}$$

where U_n is the surface of the $n-$dimensional unit sphere and $d\sigma$ is the unit spherical measure. This change of variable was shown in Arasaratnam and Haykin (2009). The integral (2.36) can therefore be rewritten as a radial integral in the form of:

$$I(f) = \int_0^\infty S(r) r^{n-1} \exp(r^2) dr, \tag{2.37}$$

where the function $S(r)$ is defined by the spherical integral given by

$$S(r) = \int_{U_n} f(rz) d\sigma(z). \tag{2.38}$$

Once the integral $I(f)$ has been decomposed into radial and spherical integrals, each integral can be solved using different rules. For the spherical integral $S(r)$, a third-degree spherical cubature rule can be used (Genz, 2003). The spherical cubature rule approximates the integral (2.38) in the form

$$\int_{U_n} f(z)d\sigma(z) \approx \omega \sum_{i=1}^{2n} f[u_i], \tag{2.39}$$

where $[u_i]$ is the i-th point of the generator u. A generator is defined as the set of points that can be obtained by permutating and changing the sign of the vector u in all possible ways. The generator is described by the following equation

$$u = (u_1, u_2, ..., u_r, 0, ..., 0) \in \mathbb{R}^n, \quad u_i \geq u_{i+1} > 0, \quad i = 1, 2, ...(r-1)$$

For example, $[1] \in \mathbb{R}^2$ constitute the set of points

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} -1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\}.$$

The values for the parameters of the cubature rule are $u^2 = 1$ and $\omega = \frac{A_n}{2n}$ in which A_n is the surface area of the unit sphere given by $A_n = \frac{2\sqrt{\pi^n}}{\Gamma(n/2)}$ where the gamma function is $\Gamma(n) = \int_o^\infty x^{n-1} \exp(-x)dx$. Once spherical integral (2.38) is solved, the integral (2.37) still needs to be solved. For solving this integral, the change of variable $t = x^2$ achieves the transformation

$$\int_0^\infty f(x)x^{n-1}exp(-x^2)dx = \frac{1}{2}\int_0^\infty \hat{f}(t)t^{\frac{n}{2}-1} \exp(-t)dt, \tag{2.40}$$

where $\hat{f}(t) = f(\sqrt{t})$. The right side of the integral (2.40) can be solved with the Gauss-Laguerre quadrature (Stroud and Secrest, 1966), which is used for solving integrals of the form

$$\int_0^\infty x^\alpha e^{(-x)} f(x) dx.$$

Using the first-degree of the generalized Gauss-Laguerre rule for $\hat{f}(t) = 1, t.$ we obtain the following equality

$$\int_0^\infty \hat{f}(t) t^{\frac{n}{2}-1} \exp(-t) dt = \omega_1 \hat{f}_i(t_1), \qquad i = 0, 1 \tag{2.41}$$

where $\hat{f}_0(t) = 1$ and $\hat{f}_0(t) = t$. Substituting the values of the functions $\hat{f}_0(t)$ and $\hat{f}_1(t)$, the following couple of equations are obtained

$$\hat{f}_0(t) = 1: \quad \omega_1 = \int_0^\infty t^{\frac{n}{2}-1} exp(-t) dt = \Gamma(\frac{n}{2})$$
$$\hat{f}_1(t) = t: \quad \omega_1 t_1 = \int_0^\infty t^{\frac{n}{2}} exp(-t) dt = \Gamma(\frac{n+1}{2})$$

solving these equations yields to $t_1 = \frac{n}{2}$ and $\omega_1 = \Gamma(\frac{n}{2})$. Therefore, the integral (2.40) is approximated as

$$\int_0^\infty f(x) x^{n-1} \exp(-x^2) dx \approx \frac{1}{2} \Gamma(\frac{n}{2}) f(\sqrt{\frac{n}{2}}). \tag{2.42}$$

Once both integrals are solved, the initial integral (2.35) can be solved using both rules as shown by

$$I(f) = \int_{\mathbb{R}^n} f(x) \exp(-x^t x) \approx \sum_{j=1}^{m_s} \sum_{i=1}^{m_r} a_i b_j f(r_i s_j), \tag{2.43}$$

where a_i and b_i are the weights obtained from the spherical cubature rule and the Gauss-Laguerre quadrature rule, and $f(r_i s_j)$ are the cubature points. Using the values of the weights and cubature points previously obtained, the third-degree spherical radial cubature rule is described by the equation

$$I(f) \approx \frac{\sqrt{\pi^n}}{2n} \sum_{i=1}^{2n} f\Big(\sqrt{\frac{n}{2}}[1]_i\Big), \tag{2.44}$$

where $[1]_i$ is an unit vector to the direction of coordinate axis.

With the use of the spherical-radial third-degree cubature rule and the Bayesian filter, the time update and measurement update of the CKF can be obtained (Arasaratnam, 2009) as follows:

2.4.1 Time Update

From the initial conditions, factorize the old posterior covariance as

$$P_{k-1|k-1} = S_{k-1|k-1} S_{k-1|k-1}^T. \tag{2.45}$$

Evaluate the cubature points ($i = 1, 2, ..., m$) as shown by

$$X_{i,k-1|k-1} = S_{k-1|k-1}\epsilon_i + \hat{x}_{k-1|k-1}. \tag{2.46}$$

Next, use the process equation to evaluate the propagated cubature points as:

$$X^*_{i,k|k-1} = f(X_{i,k-1|k-1}). \tag{2.47}$$

Once these points have been evaluated, the predicted state can be computed with the average of all the propagated points as:

$$\hat{x}_{k|k-1} = \frac{1}{m}\sum_{i=1}^{m} X^*_{i,k|k-1}. \tag{2.48}$$

The corresponding error covariance is calculated using the propagated cubature points, the predicted state, and the covariance of the process noise. This is shown as:

$$P_{k|k-1} = \frac{1}{m}\sum_{i=1}^{m} X^*_{i,k|k-1} X^{*T}_{i,k|k-1} - \hat{x}_{k|k-1}\hat{x}^T_{k|k-1} + Q_{k-1}. \tag{2.49}$$

With equation (2.49) the time update is concluded. We can now move forward to compute the measurement update.

2.4.2 Measurement Update

Factorize the predicted covariance in the form

$$P_{k|k-1} = S_{k|k-1} S^T_{k|k-1}. \tag{2.50}$$

Use the square-root of the predicted covariance and the predicted state to calculate the new cubature points $(i = 1, ..., m)$, as shown by

$$X_{i,k|k-1} = S_{k|k-1}\epsilon_i + \hat{x}_{k|k-1}. \tag{2.51}$$

The propagated cubature points are calculated using the measurement function, as shown

$$Z_{i,k|k-1} = h(X_{i,k|k-1}). \tag{2.52}$$

Using the propagated cubature points, the predicted measurement is calculated as:

$$\hat{z}_{k|k-1} = \frac{1}{m}\sum_{i=1}^{m} Z_{i,k|k-1}. \tag{2.53}$$

Using the predicted measurement in equation (2.53), the propagated cubature points (2.52), and the measurement noise covariance (R_k), compute the innovation covariance matrix as

$$P_{zz,k|k-1} = \frac{1}{m}\sum_{i=1}^{m} Z_{i,k|k-1} Z_{i,k|k-1}^T - \hat{z}_{k|k-1}\hat{z}_{k|k-1}^T + R_k. \tag{2.54}$$

The cross-covariance matrix is calculated in a similar fashion, using the cubature points (2.51), the propagated cubature points (2.52), the predicted state (2.48) and predicted measurement (2.53). This matrix is described by the following equation.

$$P_{zz,k|k-1} = \frac{1}{m}\sum_{i=1}^{m} X_{i,k|k-1} Z_{i,k|k-1}^T - \hat{x}_{k|k-1}\hat{z}_{k|k-1}^T. \tag{2.55}$$

The updated state is therefore calculated in the same form as the traditional linear Kalman filter with the equation

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + W_k(z_k - \hat{z}_{k|k-1}), \tag{2.56}$$

where the Kalman gain (W_k) is given by

$$W_k = P_{xz,k|k-1} P_{zz,k|k-1}^{-1}. \tag{2.57}$$

The corresponding covariance is obtained using the predicted covariance, the innovation covariance and the Kalman gain, as shown by

$$P_{k|k} = P_{k|k-1} - W_k P_{zz,k|k-1} W_k^T. \tag{2.58}$$

With this equation we conclude the details of the CKF.

So far we have mentioned the most relevant nonlinear Kalman filters, in which the CKF stands out as the state of the art in discrete time nonlinear filters. These filters have been widely adopted for practical applications. The approximations and assumptions required for their operation are usually accepted; however, there is another big branch of filters that are well established and have a completely different approach to state estimation. These filters are known as particle filters and are discussed in the next section.

2.5 Particle Filter (PF)

Particle filters are based on the idea of Monte Carlo integration (Doucet et al., 2001). Particle filters use a large number of weighted samples, called particles, to approximate a posterior density. Monte Carlo integration had been known for a long time but it was not used for filtering due to the degeneracy problem, that Monte Carlo methods suffer over time. The degeneracy problem occurs when only a few particles have a significant weight and all the others particles have very small weights. To solve this problem, Gordon introduced a clever method to alleviate this problem and PFs became widely popular (Gordon et al., 1993). As explained in the Bayesian filter section, the objective of an estimator

is to obtain the posterior probability density function $p(x_k|z_{1:k})$. Instead of using a time update and measurement update, PFs approximate the probability density function by using a set of samples. This approximation is described by the equation:

$$p(x_k|z_{1:k}) \approx \sum_{i=0}^{N_s} \omega_k^i \delta(x_k - x_k^i), \tag{2.59}$$

where ω_k^i refers to the weight of the ith sample at time $t = k$, and $\delta(\cdot)$ is the Dirac delta function that represents each particle. According to the principle of *importance sampling* (Arulampalam et al., 2002) the weights of the samples are given by

$$\omega_k^i = \omega_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i)|x_{k-1}^i}{q(x_k^i|x_{k-1}^i, z_k)}, \tag{2.60}$$

where $q(x_k^i|x_{k-1}^i, z_k)$ is a proposed density named the importance density. Different types of PFs have been proposed depending on the importance density selected. The common and simple choice has been to select the importance density equal to the prior, as described by the equation

$$q(x_k|x_{k-1}^i, z_k) = p(x_k|x_{k-1}^i). \tag{2.61}$$

As a result, substituting equation (2.61) in equation (2.60); yields

$$\omega_k^i = \omega_{k-1}^i p(z_k|x_k^i). \tag{2.62}$$

The result shown in equation (2.62) is one of the possible and popular implementations of particle filters; nonetheless, other possibilities exist for the important distribution. For more details please refer to Arulampalam et al., (2002).

Some advantages of particle filters are: They do not require the Gaussian assumption and can be applied to nonlinear filtering. Their coding is relatively simple. The main disadvantage of using PFs is their computational complexity. This is because a large number of particles are required by the filter, especially for problems with high state-space dimension. Additionally, no rigorous methods exist to determine the optimal number of particles required for operation; this must be determined by the designer.

In this chapter we have described the two most important types of filters, that is, approximate Kalman filters and particle filters. One the one hand, approximate Kalman filters solve the Bayesian filter theory based on Gaussian approximations and different methods to compute the integrals composed by the multiplication of nonlinear functions and Gaussian probability density functions. As a result, these filters are computationally efficient with all their parameters well defined. On the other hand, particle filters solve the Bayesian filter theory based on a large amount of particles; however, due to the lack of a Gaussian approximation and the use of particles, some of the parameters are free and have to be adjusted depending on the application. In addition, the computational time required for these filters significantly bigger compared to the approximate Kalman filters.

The discrete-time filters presented in this chapter have been traditionally applied to practical applications due to their simplicity; nevertheless, in real life, the dynamics of a system operate on a continuous-time basis. This translates into a simplification performed by the discrete-time calculation methods. To design techniques that have a better modeling of continuous-time dynamics, a

different category of filters have been developed. These algorithms recognize the continuous-time dynamics of real life applications and are known as a continuous-discrete filters. The term 'continuous' refers to the dynamics of the system and the term 'discrete' refers to the availability of the measurements, in which discrete is the true case in the majority of practical applications. Research in the continuous-discrete domains is considerably less advanced than in discrete-time counterparts; nonetheless, in the following section, I will describe the two fundamental approaches for filtering in this area.

Chapter 3

Filtering for Continuous-Discrete Systems

In most practical applications, sensors that sample time intervals are employed to obtain measurements of the state of a system. In that sense, the assumption made by discrete-time filters of receiving the measurements in a discrete-time domain is correct; however, as previously indicated, real life systems have dynamics in continuous-time. In consequence, filtering techniques that acknowledge both time domains, that is, the dynamics in continuous-time and the measurements in discrete-time, are the optimal choice for multiple applications such as tracking and navigation (Crouse, 2015; Arasaratnam et al., 2010; Kulikov and Kulikova, 2016; Åström, 2012). These filters are known as continuous-discrete filters. Conductive to the modelling of a non-linear continuous-time process, the following stochastic differential equation (SDE) is

defined for the instantaneous change of the state of the system as

$$dx(t) = f(x(t), t)dt + \sqrt{Q}d\beta(t), \tag{3.1}$$

where $x(t)$ is the $n-$dimensional state of the system at time t; $f : \mathbb{R}^n \to \mathbb{R}^n$ is a known drift function; β is the $n-$ dimensional standard Brownian process with increment $d\beta(t)$ that is independent of $x(t)$ with diffusion matrix Q.
The system is considered to be observed through noisy measurements and the relation between the measurements and the state of the system is given by the measurement equation

$$z_k = h(x_k, k) + w_k, \tag{3.2}$$

where k refers to a discrete time index, $z_k \in \mathbb{R}^m$; $h : \mathbb{R}^n \to \mathbb{R}^m$; and the measurement noise w_k is normally distributed with zero mean and covariance R_k with $R_k > 0$.
In the same fashion as discrete-time filtering, the Bayesian filter for continuous-discrete systems establishes the mathematical framework to construct the state estimate in conceptual terms (Ho and Lee, 1964). The optimal continuous-discrete Bayesian filter consists of:

- Propagation of the old posterior density between the measurement instants; known as a time update.
- Use Bayes' rule to update the posterior density; known as a measurement update.

An important difference between the time-discrete filter and continuous-discrete Bayesian filter is the time update procedure. In both cases, the measurement

update procedure is entirely equal. Given this circumstance, we next focus on describing the time update procedure.

3.1 Time Update for Continuous-Discrete Bayesian Filter

For a continuous-discrete system, the time prediction of the initial distribution of the old posterior density of the state obeys the Fokker-Planck equation (Fuller, 1969) given by

$$\frac{\partial}{\partial t}p(x(t)) = -\frac{\partial}{\partial x(t)}\Big(f(t,x(t))p(x(t))\Big) + \frac{1}{2}Q\frac{\partial^2}{\partial x(t)^2}p(x(t)), \qquad (3.3)$$

where $t < t_{t_k+1}$.

Exact solutions of the Fokker-Planck equations are only known for a few special cases. For example, with a linear process equation, it is solved by the Kalman-Bucy filter (Kalman and Bucy, 1961). Other special case is the Benes-type nonlinear process equation (Beneš, 1981). Given this circumstance, the Fokker-Planck equation has to be solved with an approximation method.

In the literature, we can find two different methods for approximating the Fokker-Planck equation. The first one is known as *linearized discretization* and the second one is known as *discretized linearization*. These approximations operate under different assumptions and because the superiority of one approach over the other has not been established, both approximations remain a focus of research.

3.2 Linearized-Discretization

This approach utilizes the Ito-Taylor expansion to discretize the SDE (3.1). As special case in Arasaratnam et al., (2010), proposed the use of the Ito-Taylor expansion of order 1.5 for the discretization of equation (3.1) with the following result:

$$x(t+\delta) = x(t) + \delta f(x(t),t) + \frac{1}{2}\delta^2 L_0 f(x(t),t) + \sqrt{Q}v + Lf(x(t),t)y, \quad (3.4)$$

in which the v and y are a correlated pair of zero mean Gaussian random variables, and the opertors L_0 and L_j are defined by

$$L_0 = \frac{\partial}{\partial t} + \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{j,p,q=1}^{n} \sqrt{Q}_{p,j}\sqrt{Q}_{q,j}\frac{\partial^2}{\partial x_p \partial \mathbf{x}_q}, \quad (3.5)$$

$$L_j = \sum_{i=1}^{n} \sqrt{Q}_{i,j}\frac{\partial}{\partial x_i}. \quad (3.6)$$

From (3.4), the discretized noise free process function f_d is defined as

$$f_d(x(t)) = x(t) + \delta f(x(t),t) + \frac{1}{2}\delta^2 L_0 f(x(t),t). \quad (3.7)$$

Once the SDE is linearized, the predicted mean $m(t+\delta)$ and predicted covariance $P(t+\delta)$ can be obtained as the expected value of the linearized function as it is shown by the equations

$$\begin{aligned} m(t+\delta) &= E[x_k|z_{1:k}] \\ &\approx E[f_d(x_k,kT) + \sqrt{Q}v + Lf(x_k,kT)y|z_{1:k}]. \end{aligned} \quad (3.8)$$

where $v \sim N(0, \delta \mathbf{I}_n); y \sim N(0, (\delta^3/3)\mathbf{I}_n)$; and $E[wy^T] = (\delta^2/2)\mathbf{I}_n$, in which $\mathbf{I}_n$ denotes the $n-$dimensional identity matrix. Because the noise terms are zero-mean Gaussian, we can simplify equation (3.8) to

$$m(t+\delta) = E[f_d(x(t))|z_{1:k}]. \tag{3.9}$$

Following the procedure in Arasaratnam et al., (2010) , the predicted covariance is given by

$$\begin{aligned} P(t+\delta) =& E[f_d(x(t))f_d^T(x(t))] - m(t+\delta)m(t+\delta)^T \\ &+ \frac{\delta^3}{3}E[(Lf(x(t),t))(Lf(x(t),t))^T] + \frac{\delta^2}{2}\sqrt{Q}E[(Lf(x(t),t)^T] \\ &+ \frac{\delta^2}{2}E[(Lf(x(t),t))]\sqrt{Q}^T + \delta Q. \end{aligned} \tag{3.10}$$

The errors in this approach are based on the discretization process utilized. Knowing this, the errors for the *linearized discretization* approach can be classified as follows:

- Type 1: Errors originating within the high-order terms ignored by the Ito-Taylor expansion of order 1.5.

Additionally to the errors type 1, there are errors caused by the approximate solutions of the integrals presented in the Bayesian filter; however, due to the fact that both approaches have these type of errors, we do not consider them a factor of choice. Once we have detailed the first approach and classified the errors presented, we can move forward into the description of the second approach, *discretized linearization*.

3.3 Discretized-Linearization

This approach was described in (Jazwinski, 1970; Sarkka, 2007), and it is obtained by taking expectations of the Fokker-Planck equation. Given a function $y = \phi(x, t)$ of the state of x, then the time evolution of the function is given by the equation

$$\frac{\partial}{\partial t}y = \frac{\partial \phi}{\partial x}\Big[f(x(t), t)dt + \sqrt{Q}d\beta\Big] + \frac{1}{2}\frac{\partial^2 \phi}{|\partial x(t)|^2}Qdt. \tag{3.11}$$

By taking the expectations of both sides we can get the following equation

$$\frac{\partial}{\partial t}E[y] = E\Big[\frac{\partial \phi}{\partial x}f(x(t), t)\Big] + \frac{1}{2}E\Big[Q\frac{\partial^2 \phi}{\partial x(t)^2}\Big]. \tag{3.12}$$

Using the functions $\phi(x) = x$ and $\phi(x) = x - m$, the following differential equations are obtained (Jazwinski, 1970):

$$\frac{dm}{dt} = E[f(x(t), t)], \tag{3.13}$$

and

$$\frac{dP}{dt} = E[(x(t) - m)f^T(x(t), t)] + E[f(x(t), t)(x(t) - m)^T] + Q. \tag{3.14}$$

Here it is relevant to mention that m and P do not characterize in general the predicted distribution because the predicted probability density function is not a Gaussian distribution. These equations determine the mean path and dispersion of equation (3.1); however, in the literature, these equations have

been utilized as the predicted state. With that said, it is important to classify and detail two other type of errors that exist with this approach:

- Type 2: When equations (3.13) and (3.14) are used to characterize the distribution, it is assumed that this distribution is a Gaussian variable; when it is not, using this assumption introduces errors into the modeling. Moreover, Gaussianity is assumed permanently during the time interval.

- Type 3: When solving equations (3.13) and (3.14), Runge-Kutta methods are utilized (Sarkka, 2007). These methods are approximate solutions for ordinary differential equations. These approximations are an additional source of error; nevertheless, this can be highly reduced with high order methods and computational time.

This approach has two different types of errors, and not all of them are under control of the designer. For example using global error control procedures in the Runge-Kutta methods can decrease Type 3 errors at the expense of computational time; however, Type 2 errors are not reduced with these procedures.

To summarize this chapter, first, we have shown the time prediction in continuous-discrete systems given by the Bayesian filter. Second, the two known approximate solutions for the time prediction have been described. Lastly, the errors presented in each type of approximation have been classified. With this concluded, we now move into the first proposal of this thesis that is a new type of continuous-discrete cubature Kalman filter given in the following chapter.

Chapter 4

Continuous-Discrete Fifth-Degree Cubature Kalman Filter

Recently, a new cubature Kalman filter has been proposed in the literature. Following the same procedure as the CKF, a new fifth-degree spherical-radial cubature rule was derived to improve the accuracy of the original CKF. This new filter was named as the fifth-degree cubature Kalman filter (5D-CKF) (Jia et al., 2013). The 5D-CKF was proposed for discrete-time systems, and the purpose of my work has been to extend this filter to the continuous-discrete theory using the two main approaches for dealing with the stochastic differential equation used for the modelling of the continuous-time dynamics.

4.1 Monomial Cubature Rules

Before describing the fifth-degree cubature rule, it is important to make clear the connotation of the degree in the context of cubature rule. Hence, the definition of a dth-degree rule is described next.

Given a vector $x = [x_1\, x_2\, ...x_n^T] \in \mathbb{R}^n$ and a weighting function $w(x)$. A dth-degree rule is defined as the approximate solution given by the equation

$$\int_{\mathbb{R}^n} f(x)w(x)dx \approx \sum_i \omega_i f(\phi_i)$$

if it is exact for $f(x)$ whose components are linear combinations of monomials $x_1^{d_1}x_2^{d_2}...x_n^{d_n}$ with d_j being nonnegative integers and $\sum_{j=1}^n d_j \leq d$. This indicate that the higher the degree of the cubature rule, the more accurate its solution becomes.

4.2 Fifth-Degree Cubature Rule

To obtain the fifth-degree cubature rule, we recall the radial transformation seen in Chapter 2 given by the equations

$$I(f) = \int_{\mathbb{R}^n} f(x)\exp(-x^T x)dx, \tag{4.1}$$

where the vector $x \in \mathbb{R}^n$. Applying the change of variables from the Cartesian to the radial vector r, let $x = rz$ with $z^T z = 1$ such that $x^T x = r^2$ for $r \in [0, \infty)$

the integral (4.1) becomes

$$I(f) = \int_0^\infty \int_{U_n} f(rz) r^{n-1} \exp(-r^2) d\sigma(z) dr, \tag{4.2}$$

and the integral (4.2) can be approximated as shown by

$$I(f) = \int_{\mathbb{R}^n} f(x) \exp(-x^t x) \approx \sum_{j=1}^{m_s} \sum_{i=1}^{m_r} a_i b_j f(r_i s_j), \tag{4.3}$$

where a_i are the weights obtained from the spherical cubature rule; b_i are the weights obtained from the the radial rule, and $f(r_i s_j)$ are the cubature points. To create the new cubature rule, a higher degree spherical and a radial cubature rules were utilized. The new symmetric spherical rule establishes the following approximation (Genz, 2003). For the spherical integral $\int_{U_n} f(Z) d\sigma$, the approximation given by

$$\int_{U_n} f(Z) d\sigma \approx \sum_{|p|=m} \omega_p f\{U_p\}, \tag{4.4}$$

is a $(2m+1)$th-degree rule to approximate the integral, where the terms $Z = (z_1, z_2, ..., z_n)$, and $U_n = \{Z | Z \in \mathbb{R}^n, z_1^2 + z_2^2 + ... + z_n^2 = 1\}$ and σ is an element of surface on U_n. The weights are defined as

$$\omega_p = I_{u_n} \left(\sum_{|p|=m} \prod_{i=1}^{n} \prod_{j=0}^{p_i - 1} \frac{z_i^2 - u_j^2}{u_{pi}^2 - u_j^2} \right), \tag{4.5}$$

and the cubature points are given by

$$f\{U_p\} = 2^{-c(U)} \sum_s f(s_1 u_1, s_2 u_2, ..., s_n u_n), \tag{4.6}$$

where $u_i = \sqrt{i/m}$; $c(U)$ is the number of nonzero entries in $(u_1, u_2, ..., u_n)$, and the sum $\sum_s$ is taken over all of the sign combinations that occur when $s_i = \pm 1$. The subscripts p_i are nonnegative integers with $p = [p_1, p_2, ..., p_n]$ and $|p| = p_1 + p_2 + ... + p_n = m$. The points $[s_1 u_{p_1}, s_2 u_{p_2}, ..., s_n u_{p_n}]^T$ have the associated weitghts $2^{-c(Up)}\omega_p$. Using the value of $m = 2$ for a fifth-degree spherical rule, the following rule with $2n^2$ weights is obtained

$$\begin{aligned}\int_{U_n} f(z)d\sigma(z) \approx \omega_1 \sum_{j=1}^{n(n-1)} \Big(f(s_j^+) + f(-s_j^+) + f(s_j^-) + f(-s_j^-) \Big)) \\ + \omega_2 \sum_{j=1}^{n} (f(e_j) + f(-e_j)),\end{aligned} \tag{4.7}$$

where the weights ω_1 and ω_2 are given by:

$$\omega_1 = \frac{A_n}{n(n+2)}, \tag{4.8}$$

and

$$\omega_2 = \frac{(4-n)A_n}{2n(n+2)}. \tag{4.9}$$

The set of points s_j^+ and s_j^- are defined as:

$$\{s_j^+\} \triangleq \Big\{ \sqrt{\frac{1}{2}}(e_k + e_l) : k < l, k, l = 1, 2, ..., n \Big\}, \tag{4.10}$$

and

$$\{s_j^-\} \triangleq \Big\{ \sqrt{\frac{1}{2}}(e_k - e_l) : k < l, k, l = 1, 2, ..., n \Big\}, \tag{4.11}$$

where e_i is a unit vector to the direction of coordinate axis i. Once the spherical rule has been obtained, a radial rule of the same degree also must be obtained. The radial rule utilized is calculated based on the moment matching method.

The objective of the method is to find the points r_i and weights $w_{r,i}$ to satisfy the equation:

$$\int_0^\infty S(r) r^{n-1} \exp(r^2) dr = \sum_{i=1}^{N_r} \omega_{r,i} S(r_i), \tag{4.12}$$

where $S(r) = r^l$ is a monomial in r with l being the order of the monomial. To construct a $(2m+1)$th-degree radial rule, equation (4.12) needs to be exact for $(l = 0, 2, ..., 2m)$ which leads to $(m+1)$ equations. Knowing that the solution to the integral on the right side of equation (4.12) is equal to $\frac{1}{2}\Gamma(\frac{n+l}{2})$ for $l = 0, 2, ..., 2m$, a fifth-degree radial rule can be constructed from the following equations:

$$S(r) = r^0 : \qquad \omega_{r,1} r_1^0 + \omega_{r,2} r_2^0 = \frac{1}{2}\Gamma\Big(\frac{1}{2}\Big), \tag{4.13}$$

$$S(r) = r^2 : \qquad \omega_{r,1} r_1^2 + \omega_{r,2} r_2^2 = \frac{1}{2}\Gamma\Big(\frac{1}{2}+1\Big), \tag{4.14}$$

and

$$S(r) = r^4 : \qquad \omega_{r,1} r_1^4 + \omega_{r,2} r_2^4 = \frac{1}{2}\Gamma\Big(\frac{1}{2}+2\Big). \tag{4.15}$$

Because there are only three equations and four variables, we can set $r_1 = 0$. Solving the system of equations, we now have the following points

$$\begin{aligned} r_1 &= 0, \\ r_2 &= \sqrt{\frac{1}{2}n + 1}, \end{aligned} \tag{4.16}$$

and the weights are given by

$$\begin{aligned} \omega_{r,1} &= \frac{1}{(n+2}\Gamma(\frac{1}{2}n), \quad \text{and} \\ \omega_{r,2} &= \frac{n}{2(n+2)}\Gamma(\frac{1}{2}n). \end{aligned} \tag{4.17}$$

Combining the fifth-degree spherical rule and the fifth-degree radial rule in exactly the same fashion as it was done for the CKF, the fifth-degree spherical-radial cubature rule is obtained as:

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)N(x;0,I) \approx \\ &\frac{2}{n+2}f(0) + \frac{1}{(n+2)^2}\sum_{i=1}^{n(n-1)/2}\Big(f(\sqrt{n+2}\cdot s_j^+) + f(\sqrt{n+2}\cdot s_j^+\Big) \\ &+ \frac{1}{(n+2)^2}\sum_{i=1}^{n(n-1)/2}\Big(f(\sqrt{n+2}\cdot s_j^-) + f(\sqrt{n+2}\cdot s_j^-)\Big) \\ &+ \frac{4-n}{2(n+2)^2}\sum_{i=1}^{n}\Big(f(\sqrt{n+2}\cdot e_i) + f(\sqrt{n+2}\cdot e_i)\Big).\end{aligned} \tag{4.18}$$

Once the fifth-degree cubature rule have been completely described, we can use the approaches mentioned in Chapter 3 to develop the continuous-discrete fifth-degree cubature Kalman filter (CD-5D-CKF).

4.3 Continuous-Discrete Fifth-Degree Cubature Kalman Filter (CD-5D-CKF)

4.3.1 Linearized-Discretization

The first approach used for the CD-5D-CKF is the *linearized discretization*. For this purpose, it is necessary to review the equations in Chapter 3. The

predicted state is given by

$$\begin{aligned} m(t+\delta) &= E[f_d(x_k, kT)|z_{1:k}] \\ &= \int_{\mathbb{R}^n} f_d(x_k, kT) N(x_k; \hat{x}_k|k, P_{k|k}) dx_k, \end{aligned} \tag{4.19}$$

and the predicted covariance is defined as

$$\begin{aligned} P(t+\delta) =& E[f_d(x(t)) f_d^T(x(t))] - m(t+\delta)m(t+\delta)^T + \frac{\delta^3}{3}(Lf(x(t),t))(Lf(x(t),t))^T \\ &+ \frac{\delta^2}{2}\sqrt{Q}(Lf(x(t),t)^T + \frac{\delta^2}{2}(Lf(x(t),t))\sqrt{Q}^T + \delta Q. \end{aligned} \tag{4.20}$$

Using the definition of expected value of a random variable in equation (4.20), we obtain the following result for the predicted covariance

$$\begin{aligned} P(t+\delta) = & \int_{\mathbb{R}^\kappa} f_d(x_k, kT) f_d^T(x_k, kT) N(x_k; \hat{x}_{k|k, P_{k|k}}) dx_k \\ & - m(t+\delta)m(t+\delta)^T + \frac{\delta^3}{3}(Lf(x(t),t))(Lf(x(t),t))^T \\ & + \frac{\delta^2}{2}\sqrt{Q}(Lf(x(t),t)^T + \frac{\delta^2}{2}(Lf(x(t),t))\sqrt{Q}^T + \delta Q. \end{aligned} \tag{4.21}$$

Using the weights and cubature points of the fifth-degree cubature rule to solve the integrals in equations (4.19) and (4.21), the predicted state and the

predicted covariance are given by the equations (4.22) and (4.23) respectively.

$$
\begin{aligned}
m^{j+1}_{k-1|k-1} = \frac{1}{(n+2)^2} \sum_{j=1}^{n(n-1)/2} \Big(& f_d(\sqrt{n+2}\sqrt{P_{k-1}}s^+_j + m^j_{k-1|k-1}, kT + j\delta) \\
& + f_d(-\sqrt{n+2}\sqrt{P_{k-1}}s^+_j + m^j_{k-1|k-1}, kT + j\delta)\Big) \\
+ \frac{1}{(n+2)^2} \sum_{j=1}^{n(n-1)/2} \Big(& f_d(\sqrt{n+2}\sqrt{P_{k-1}}s^-_j + m^j_{k-1|k-1}, kT + j\delta) \\
& + f_d(-\sqrt{n+2}\sqrt{P_{k-1}}s^-_j + m^j_{k-1|k-1}, kT + j\delta)\Big) \\
+ \frac{4-n}{2(n+2)^2} \sum_{i=1}^{n} \Big(& f_d(\sqrt{n+2}\sqrt{P_{k-1}}e_i + m^j_{k-1|k-1}, kT + j\delta) \\
& + f_d(-\sqrt{n+2}\sqrt{P_{k-1}}e_i + m^j_{k-1|k-1}, kT + j\delta)\Big) \\
+ \frac{2}{n+2} f_d(m^j_{k-1|k-1}),
\end{aligned}
\tag{4.22}
$$

$$
\begin{aligned}
& P^{j+1}_{k-1|k-1} = \\
& \frac{1}{(n+2)^2}\left[\left(N^{j+1}_{k-1|k-1}(s^+_j)\right)\left(N^{j+1}_{k-1|k-1}(s^+_j)\right)^T + \left(N^{j+1}_{k-1|k-1}(-s^+_j)\right)\left(N^{j+1}_{k-1|k-1}(-s^+_j)\right)^T\right] \\
+ & \frac{1}{(n+2)^2}\left[\left(N^{j+1}_{k-1|k-1}(s^-_j)\right)\left(N^{j+1}_{k-1|k-1}(s^-_j)\right)^T + \left(N^{j+1}_{k-1|k-1}(-s^-_j)\right)\left(N^{j+1}_{k-1|k-1}(-s^-_j)\right)^T\right] \\
+ & \frac{4-n}{2(n+2)^2}\left[\left(N^{j+1}_{k-1|k-1}(e_i)\right)\left(N^{j+1}_{k-1|k-1}(e_i)\right)^T + \left(N^{j+1}_{k-1|k-1}(-e_i)\right)\left(N^{j+1}_{k-1|k-1}(-e_i)\right)^T\right] \\
+ & \frac{\delta^2}{2}\left(Lf(m^j_{k-1|k-1})Q^T + QLf(m^j_{k-1|k-1})^T\right) + \frac{\delta^3}{3}\left(Lf(m^j_{k-1|k-1})Lf(m^j_{k-1|k-1})^T\right) \\
+ & \delta Q,
\end{aligned}
\tag{4.23}
$$

where the vector $N^{(j+1)}_{k|k}(\Phi)$ is defined by

$$N_{k|k}^{(j+1)}(\Phi) = \Big[N_{1,k|k}^{j+1}(\Phi) - m_{k-1|k-1}^{j+1} \quad N_{2,k|k}^{j+1}(\Phi) - m_{k-1|k-1}^{j+1} \quad ... N_{N,k|k}^{j+1}(\Phi) - m_{k-1|k-1}^{j+1}\Big], \tag{4.24}$$

and each element in the vector (4.24) is computed by

$$N_{i,k-1|k-1}^{j+1}(\Phi) = f_d\Big(m_{k-1|k-1}^{j} + \sqrt{P_{k-1}}\sqrt{n+2}\,\Phi, kT + j\delta\Big), \tag{4.25}$$

where $s_j^+$and s_j^- are given in equations (4.10) and (4.11), respectively, and f_d is the noise free function given in equation (3.7). This concludes the description of the *linearized discretization* approach. We now move into the description of the *discretized linearization* approach.

4.3.2 Discretized-Linearization

The time prediction of the CD-5D-CKF based on the *discretized linearization* approach can be obtained by substituting the fifth-degree cubature rule into the predicted state (3.13) and predicted covariance (3.14) equations discussed in Chapter 3, and this results in the following:

$$\frac{dm}{dt} = E[f(x(t), t)], \tag{4.26}$$

$$\frac{dP}{dt} = E[(x(t) - m)f^T(x(t), t)] + E[f(x(t), t)(x(t) - m)^T] + Q. \tag{4.27}$$

Therefore, the predicted state and predicted covariance of the CD-5D-CKF are obtained by the integration of the following differential equations from the initial conditions m_{k-1}, P_{k-1} at time t_{k-1} to t_k.

$$
\begin{aligned}
&\frac{dP(t)}{dt} = \\
&\frac{1}{(n+2)^2}\sum_{i=1}^{n(n-1)/2}\Big[f(m(t)+\sqrt{P(t)}\cdot\sqrt{n+2}\cdot s_j^+)\cdot s_j^{+T}\cdot\sqrt{n+2}\cdot\sqrt{P(t)}^T \\
&\qquad + f(m(t)-\sqrt{P(t)}\cdot\sqrt{n+2}\cdot s_j^+)\cdot -s_j^{+T}\cdot\sqrt{n+2}\cdot\sqrt{P(t)}^T \\
&\qquad + s_j^+\cdot\sqrt{n+2}\cdot\sqrt{P(t)}\cdot f^T(m(t)+\sqrt{P(t)}\cdot\sqrt{n+2}\cdot s_j^+) \\
&\qquad + -s_j^+\cdot\sqrt{n+2}\cdot\sqrt{P(t)}\cdot f^T(m(t)-\sqrt{P(t)}\cdot\sqrt{n+2}\cdot s_j^+)\Big] \\
&+\frac{1}{(n+2)^2}\sum_{i=1}^{n(n-1)/2}\Big[f(m(t)+\sqrt{P(t)}\cdot\sqrt{n+2}\cdot s_j^-)\cdot s_j^{-T}\cdot\sqrt{n+2}\sqrt{P(t)}^T \\
&\qquad + f(m(t)-\sqrt{P(t)}\cdot\sqrt{n+2}\cdot s_j^-)\cdot -s_j^{-T}\cdot\sqrt{n+2}\sqrt{P(t)}^T \\
&\qquad + s_j^-\cdot\sqrt{n+2}\cdot\sqrt{P(t)}\cdot f^T(m(t)+\sqrt{P(t)}\cdot\sqrt{n+2}\cdot s_j^-) \\
&\qquad + -s_j^-\cdot\sqrt{n+2}\cdot\sqrt{P(t)}\cdot f^T(m(t)-\sqrt{P(t)}\cdot\sqrt{n+2}\cdot s_j^-)\Big] \\
&+\frac{4-n}{2(n+2)^2}\sum_{i=1}^{n}\Big[f(m(t)+\sqrt{P(t)}\cdot\sqrt{n+2}\cdot e_i)\cdot e_i^T\cdot\sqrt{n+2}\sqrt{P(t)}^T \\
&\qquad + f(m(t)-\sqrt{P(t)}\cdot\sqrt{n+2}\cdot e_i)\cdot -e_i^T\cdot\sqrt{n+2}\sqrt{P(t)}^T \\
&\qquad \cdot + e_i\cdot\sqrt{n+2}\sqrt{P(t)}\cdot f^T(m(t)+\sqrt{P(t)}\cdot\sqrt{n+2}\cdot e_i) \\
&\qquad \cdot + - e_i\cdot\sqrt{n+2}\sqrt{P(t)}\cdot f^T(m(t)-\sqrt{P(t)}\cdot\sqrt{n+2}\cdot e_i)\Big],
\end{aligned}
\tag{4.28}
$$

and

$$\frac{dm(t)}{dt} =$$
$$\frac{1}{(n+2)^2} \sum_{i=1}^{n(n-1)/2} \left(f(m(t) + \sqrt{P(t)} \cdot \sqrt{n+2} \cdot s_j^+) + f(m(t) - \sqrt{P(t)} \cdot \sqrt{n+2} \cdot s_j^+ \right)$$
$$+ \frac{1}{(n+2)^2} \sum_{i=1}^{n(n-1)/2} \left(f(m(t) + \sqrt{P(t)} \cdot \sqrt{n+2} \cdot s_j^-) + f(m(t) - \sqrt{P(t)} \cdot \sqrt{n+2} \cdot s_j^- \right)$$
$$+ \frac{4-n}{2(n+2)^2} \sum_{i=1}^{n} \left(f(m(t) + \sqrt{P(t)} \cdot \sqrt{n+2} \cdot e_i) + f(m(t) - \sqrt{P(t)} \cdot \sqrt{n+2} \cdot e_i \right)$$
$$+ \frac{2}{n+2} f(m(t)). \tag{4.29}$$

Equations (4.28) and (4.29) can be integrated with methods used for ordinarial differential equations such as the Runga-Kutta methods (Dormand and Prince, 1980). Once the time prediction is computed, the measurement update is computed identically to the discrete-time 5D-CKF described in (Jia et al., 2013). This description concludes the time prediction of the *discretized linearization* approach of the CD-5D-CKF. We now move onto the experimental analysis.

4.4 Numerical Simulations and Discussion

With the purpose of evaluating the two different implementations of the CD-5D-CKF, I have chosen to test them with the air-traffic-control scenario from Arasaratnam et al., (2010). The objective of the air-traffic-control scenario is to track a 3-dimensional target that is moving in a coordinated turn in the horizontal plane at a constant speed and turn rate as it is shown in Figure 4.1.

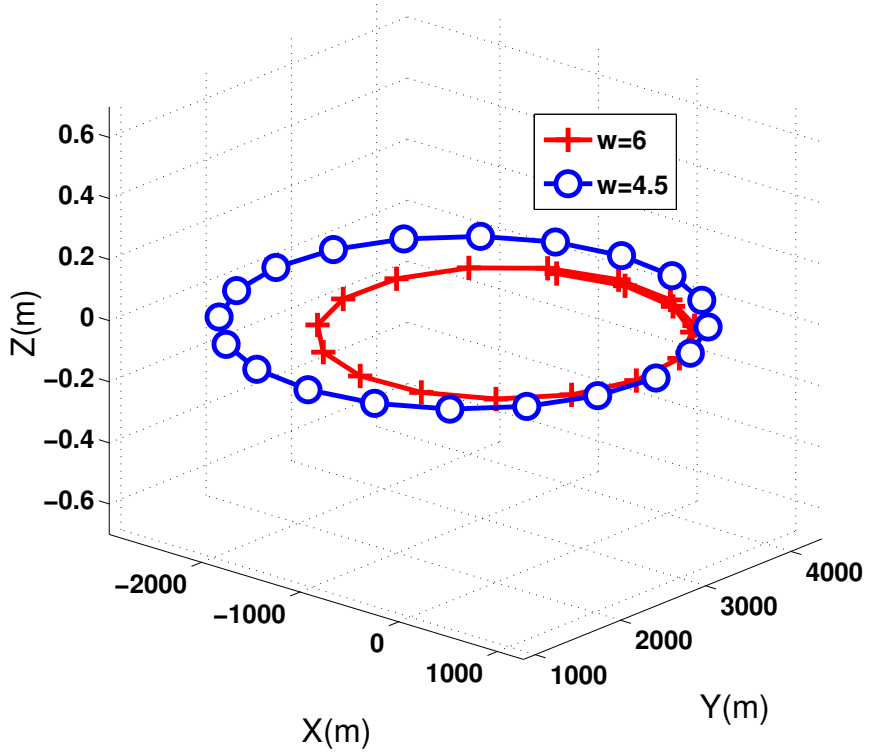

Figure 4.1: Trajectories with different turn rate $w = 4.5°/s$ and $w = 6°/s$. The trajectory with $w = 4.5°/s$ is only for comparison.

This scenario is challenging and is used for testing due to the following reasons:

- It is a practical problem without a final solution achieved.
- The problem involves nonlinear processes and nonlinear measurement models.
- It is a seven-dimensional nonlinear problem, which could possibly be the higest dimension in a tracking problem.

- There is a control over the degree of nonlinearity via the turn rate (parameter w). As this value increases, the problem becomes more difficult.
- A challenging scenario can provide clear evidence of the performance of the filters compared to experiments with few nonlinearities where the results could be almost identical.
- This scenario has been used by multiple researches for the analysis of filters (Särkkä, 2013) (Kulikov and Kulikova, 2016).

The dynamics of the aircraft obeys the stochastic differential equation given by

$$dx(t) = f(x(t))dt + \sqrt{Q}d\beta(t) \tag{4.30}$$

where the state vector is given by $x = [x(t)\ \dot{x}(t)\ y(t)\ \dot{y}(t)\ z(t)\ \dot{z}(t)\ w(t)] \in \mathbb{R}^7$ with $x(t)$, $y(t)$ and $z(t)$ denoting instantaneous positions; $\dot{x}(t)$, $\dot{y}(t)$ and $\dot{z}(t)$ denoting instantaneous velocities and $w(t)$ denotes the turn rate. The parameters that define the SDE (4.30) are the drift function given by $f(x) = [\dot{x}(t)\ \ -w(t)\dot{y}(t)\ \dot{y}(t)\ w(t)\dot{x}(t)\ \dot{z}(t)\ 0\ 0] \in \mathbb{R}^7$ and the noise term $\beta(t) = [\beta_1(t)\ \beta_2(t)\ ...\ \beta_7]^T$ with $\{\beta_i(t)\} = 1, 2, ..., 7$, being mutually independent Brownian process with zero mean and unit covariance.

The true trajectory of the object is simulated applying the Ito-Taylor transform of order 1.5 to the SDE (4.30) that gets the stochastic difference equation

$$x(t_k + (j+1)\delta) = f_d(x(t_k + j\delta)) + \sqrt{Q}v + Lf(x(t_k + j\delta)) \tag{4.31}$$

where $1 \leq j \leq m$ and $\delta = T/m$ with m=1000 time-steps/sampling interval. The sampling intervals for the experiment are T= 4, 6, 8 seconds. The driving noise diffusion matrix is $Q = \text{diag}[0, \sigma_1^2, 0, \sigma_1^2, 0, \sigma_1^2, \sigma_2^2]$ with $\sigma_1 = \sqrt{0.2}m/s$ and $\sigma_2 = 0.007°/s$. The discretized noise free process function $\mathbf{f}_d$ and the function Lf in equation (4.31) are computed as follows:

$$\mathbf{f}_d = \begin{bmatrix} x + \delta\dot{x} - \frac{\delta^2}{2}\omega\dot{(y)} \\ \dot{x} - \delta\omega\dot{y} - \frac{\delta^2}{2}\omega^2\dot{x} \\ y + \delta\dot{y} + \frac{\delta^2}{2}\omega\dot{x} \\ \dot{y} + \delta\omega\dot{x} - \frac{\delta^2}{2}\omega^2\dot{y} \\ z + \delta\dot{z} \\ \dot{z} \\ \omega \end{bmatrix}, \text{ and } \quad Lf = \begin{bmatrix} 0 & \sigma_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sigma_1\omega & 0 & 0 & \sigma_2\dot{y} \\ 0 & 0 & 0 & \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_1\omega & 0 & 0 & 0 & 0 & -\sigma_2\dot{y} \\ 0 & 0 & 0 & 0 & 0 & \sigma_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

For the computer experiment, the radar is located at the origin and the measurements received are range (r_k), azimuth angle (θ_k), and elevation angle (ρ_k). The measurement function is given by

$$\begin{bmatrix} r_k \\ \theta_k \\ \rho_k \end{bmatrix} = \begin{bmatrix} \sqrt{(x_k^2 + y_k^2 + z_k^2)} \\ tan^{-1}(y_k/x_k) \\ tan^{-1}(z_k/\sqrt{x_k^2 + y_k^2}) \end{bmatrix} + w_k,$$

where w_k is the measurement noise and $w_k \backsim N(0, R)$ with $R = diag[\sigma_r^2, \sigma_\theta^2, \sigma_\rho^2]$, where $\sigma_r = 50m$, $\sigma_\theta = 0.1°$ and $\sigma_\rho = 0.1°$. The initial state is set to be a random variable as recommended in section 5.5.2 of Bar-Shalom et al., (2004). The accumulative root-mean square error (RMSE) of the position, velocity and

turn rate were used for comparing the two different implementations of the filters. The RMSE equations for these measures are:

$$RMSE_p = \sqrt{\frac{1}{NK}\sum_{n=1}^{N}\sum_{k=1}^{K}\left(\left(x_k^n - \hat{x}_k^n\right)^2 + \left(y_k^n - \hat{y}_k^n\right)^2 + \left(z_k^n - \hat{z}_k^n\right)^2\right)}, \quad (4.32)$$

$$RMSE_v = \sqrt{\frac{1}{NK}\sum_{n=1}^{N}\sum_{k=1}^{K}\left(\left(\dot{x}_k^n - \hat{\dot{x}}_k^n\right)^2 + \left(\dot{y}_k^n - \hat{\dot{y}}_k^n\right)^2 + \left(\dot{z}_k^n - \hat{\dot{z}}_k^n\right)^2\right)}, \quad (4.33)$$

and

$$RMSE_w = \sqrt{\frac{1}{NK}\sum_{n=1}^{N}\sum_{k=1}^{K}\left(\left(w_k^n - \hat{w}_k^n\right)^2\right)}, \quad (4.34)$$

where $(x_k^n, \dot{x}_k^n, y_k^n, \dot{y}_k^n, z_k^n, \dot{z}_k^n, w_k^n)$ are the true values of the object and $(\hat{x}_k^n, \hat{\dot{x}}_k^n, \hat{y}_k^n, \hat{\dot{y}}_k^n, \hat{z}_k^n, \hat{\dot{z}}_k^n)$ are the estimated values at the time index k in the nth Monte Carlo run. The usual number of Monte Carlo simulations used in target tracking experiments has been 100, which is the case for (Arasaratnam et al., 2010; Särkkä and Solin, 2012; Kulikov and Kulikova, 2016). This is because it has been shown that in target tracking the normalize estimation error square region becomes small after this number of Monte Carlo simulations (Lin et al., 2005). Considering this, and the computational time required for the experiment, we decided to perform 1000 Monte Carlo simulations. The accumulative RMSE was obtained for the interval 0 -210 seconds.

The turn rate utilized was $w = 6°/s$. Divergence is assumed when the position error is bigger than 500 meters, that is, when the filter fails on its operation and it shown in the following inequality.

$$\sqrt{((\left(x_k^n - \hat{x}_k^n\right)^2 + \left(y_k^n - \hat{y}_k^n\right)^2 + \left(z_k^n - \hat{z}_k^n\right)} > 500m. \tag{4.35}$$

For the purpose of comparison, we included the continuous-discrete third-degree CKF implemented with the two approaches mentioned in Chapter 3. Therefore, the filtering approaches I compared were:

(1) 3D-IT-CKF: The square-root continuous-discrete third-degree CKF based in the Ito-Taylor expansion of order 1.5.

(2) 3D-RK-CKF: The square-root continuous-discrete third-degree CKF based on the discretize-linearization approach.

(3) 5D-IT-CKF: The continuous-discrete fifth-degree CKF based on the Ito-Taylor expansion of order 1.5.

(4) 5D-RK-CKF: The continuous-discrete fifth-degree CKF based on the discretize-linearization approach.

Square-root filters have been proposed in the literature to avoid operations in the covariance matrix that can eliminate working properties such as positive definiteness and symmetry. This is achieve by propagating the square-root of the covariance matrix instead of the covariance matrix, hence the name. These filters are usually more stable than their traditional implementations. Unfortunately the design of square-root filters is not trivial and there is not square-root implementation of the fifth-degree cubature Kalman filter. We con-

sider important to use these filters for the comparison because these algorithms are available in the literature

The parameter m determines the number of iterations per sampling time. For the filters based on the Ito-Taylor expansion the parameter m is selected by the designer, whereas for the filters based on the *discretized linearization* approach, the number of iterations per sampling time this is automatically determined by the Matlab Solver ODE45 that has implemented an explicit Runga-Kutta method (Shampine and Reichelt, 1997) with local error control set to $10^{-}9$. Control errors methods is a field of research for approximate methods such as Runga-Kutta methods. The objective of the error control is to guaranteed that the approximation used has an error that is minor to a desired threshold (e.g. $10^{-}9$) (Kulikov, 2013).

4.4.1 Results

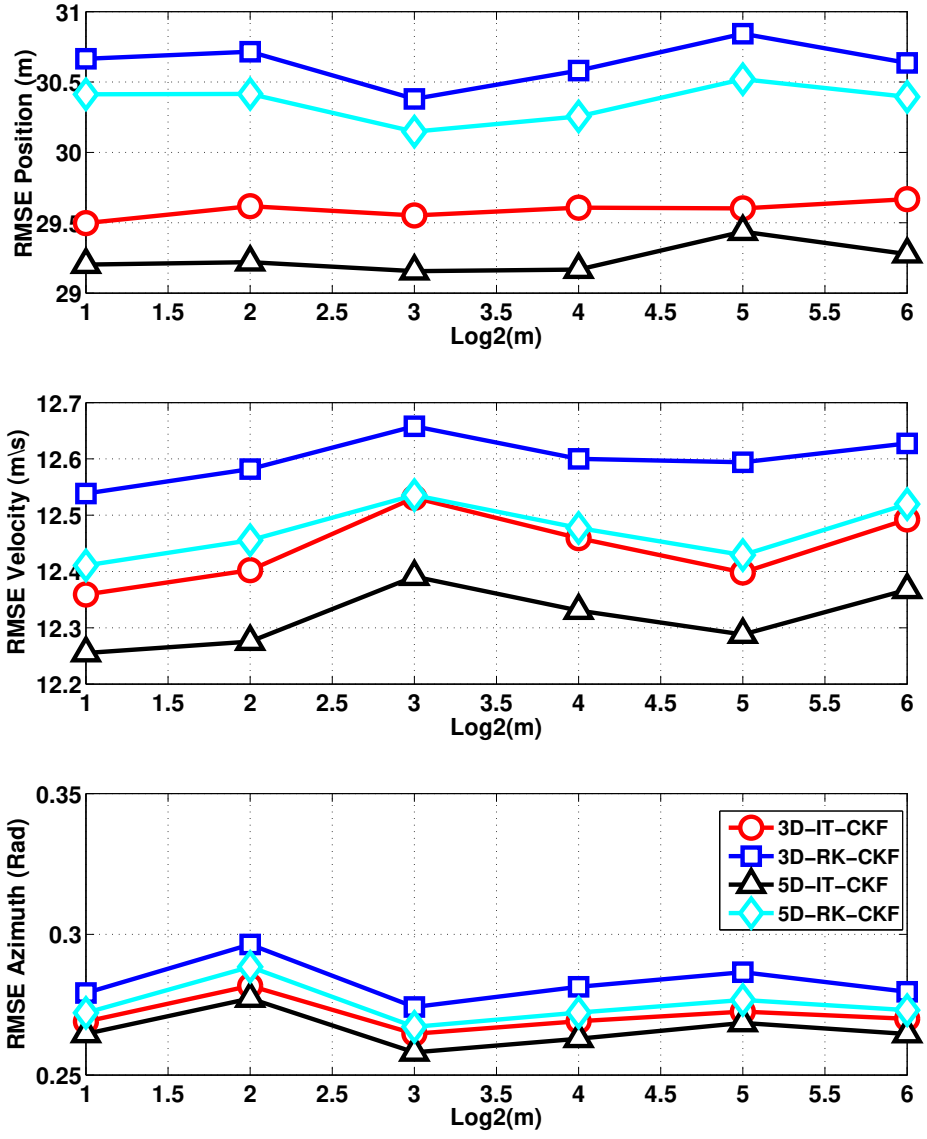

Figure 4.2: Accumulative RMSE in position, velocity and azimuth vs $log_2(m)$, where m is the number of iterations/sampling interval. Sampling time used T=4 s.

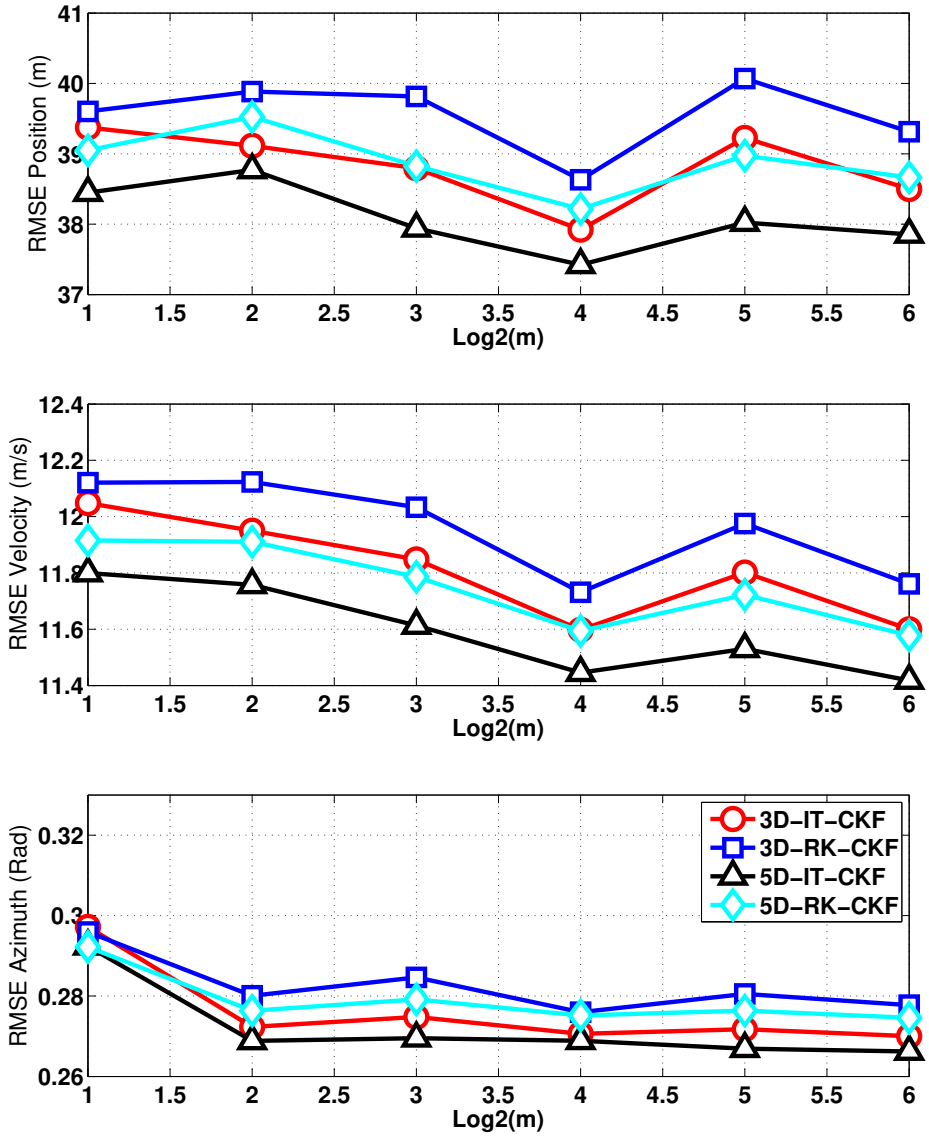

Figure 4.3: Accumulative RMSE in position, velocity and azimuth vs $log_2(m)$, where m is the number of iterations/sampling interval. Sampling time used T=6 s.

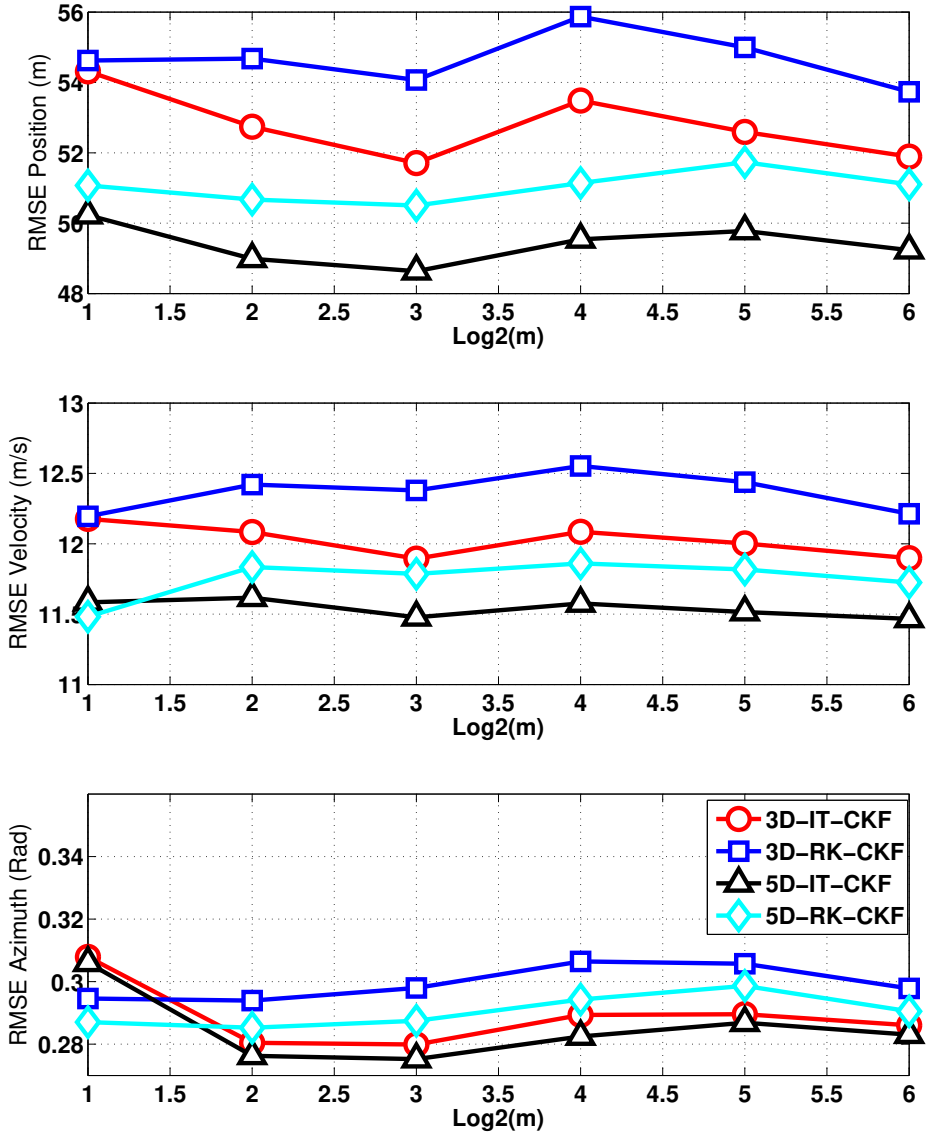

Figure 4.4: Accumulative RMSE in position, velocity and azimuth vs $log_2(m)$, where m is the number of iterations/sampling interval. Sampling time used T=8 s.

In Figure 4.2, for the sampling time set to $T = 4s$, the filters based on the Ito-Taylor expansion had better performance than their counterparts based on Runga-Kutta methods. The 5D-IT-CKF had the best performance of all of them and unexpectedly the 3D-IT-CKF outperformed the 5D-RK-CKF; however the difference in performance between the 5D-IT-CKF and the 3D-IT-CKF is very small with a difference of only 0.5 meters in position. The small variations in performance from the filters based on Runga-Kutta methods is attributed to the initial conditions; however, those initial conditions were applied in all the filters, that is the reason for the similarity in their behaviour. For the filters based on the Ito-Taylor expansion, the value of m do not seem to affect the RMSE in position, but the RMSE in azimuth decreased after the value of two, which indicates that using the value of $m = 2^3$ is convenient; bigger values of m seem not to improve the performance. The number of divergences presented for this time sample is shown in Table 4.1; considering that 6000 Monte Carlo simulations were performed for this sampling time, the number of divergences presented in the filters is minor and it can be said that all filters are equally good in that regard.

Figure 4.3 was obtained by fixing the sampling interval $T = 6s$ and the average RMSE of 1000 Monte Carlo simulations for every value of m. This scenario continues to demonstrate the superiority of the 5D-IT-CKF compared to the other filters. In a different case than the previous one, the 3D-IT-CKF did not outperformed the 5D-RK-CKF and both filters had equally good RMSE in position and velocity with a small difference in azimuth. The less accurate filter was the 3D-RK-CKF. The number of divergences for this sampling time shown in Table 4.1 exhibit that all filters had approximately the same number

of divergences. For the filters based on the Ito-Taylor expansion, it is observed that once the number of iterations has achieved the value of $m = 2^4$, there is no gain in accuracy when this value increases. For the filters based on Runga-Kutta methods, the number of iterations per sampling is calculated by the ODE solver; therefore, the variations of performance is again attributed to the initial conditions, which is consistent in the four filters compared.

The RMSE in position, velocity and azimuth for the sampling time $T = 8s$ is shown in Figure 4.4. Due to the large sampling time, this scenario is the most difficult compared to the previous ones. For such conditions, the filters based on the fifth-degree cubature rule achieved the best accuracy in position and velocity, being the 5D-IT-CKF the best of all. Among the filters based on the third-degree cubature rule, the 3D-IT-CKF outperformed the 3D-RK-CKF; which is a consistent result among the different setups used. In addition, the filters based on the Ito-Taylor expansion do not show an improvement in accuracy after the value of $m = 3$. Unfortunately, the number of divergences of the filters based on the fifth-degree cubature rule was significantly bigger compared to the filters using the third-degree cubature rule, which is shown in Table 4.1. This is explained due to the lack of a square-root implementation, which is available for the filters based on the third-degree cubature rule but not for the ones based on the fifth-degree rule. Considering this, the 3D-IT-CKF or the 3D-RK-CKF are more reliable for larger sampling times.

4.4.2 Divergences

In addition to RMSE, the number of divergences is a good measure of the stability of a filter. A filter can stop working when the covariance matrix loses its symmetry or positive semi-definiteness. For that reason, it is common that filtering algorithms try to avoid the use of operations that can produce the loss of these properties, such as the square-root filters. The comparison of a number of divergences in the experimental results is shown in Table 4.1.

Table 4.1: Number of divergences in 6000 Monte Carlo simulations for the four different types of CKFs with different sampling times T given in seconds.

Filter	$T = 4$ s	$T = 6$ s	$T = 8$ s
3D-IT-CKF	5	42	288
3D-RK-CKF	2	42	282
5D-IT-CKF	3	41	698
5D-RK-CKF	3	40	645

The results in Table 4.1 give evidence that for the sampling times $T = 4$ s and $T = 6$ s, the number of divergences of the filters based on the fifth-degree cubature rule is equally good than the filter using the third-degree cubature rules which were implemented in square-root form. Especially, considering a large number of Monte Carlo simulations; however, for the sampling time $T = 8s$, the lack of a square-root implementation was visible and the filters based on the fifth-degree cubature rule suffer from a significantly large number of divergences that their counterparts. Theses result indicates that the 5D-IT-CKF and the 5D-RK-CKF are not as stable as the 3D-RK-CKF and the 3D-IT-CKF for large sampling times.

4.4.3 Computational Time

For practical applications in filtering, the computational time has always been an issue of concern. This is because most of the time, filtering has to operate on a real time; therefore, algorithms cannot be computationally too demanding. Considering this, we included in Table 4.2 the time consumption of the filters that were analyzed.

Table 4.2: Computational time in seconds for the filters in one run using a MacBook Pro with a 2.7 GHz Intel i5 and 16GB DDR3

	Log2 (m)					
Filter	1	2	3	4	5	6
3D-IT-CKF	0.035	0.038	0.050	0.078	0.129	0.335
3D-RK-CKF	1.30	1.27	1.27	1.29	1.33	1.35
5D-IT-CKF	0.103	0.140	0.231	0.417	0.783	1.574
5D-RK-CKF	3.568	3.577	3.629	3.774	3.784	3.792

The results shown in Table 4.2 provide evidence of the faster computational time of filters that use Ito-Taylor expansion compared to their counterparts. The processing time of the 3D-IT-CKF and 5D-IT-CKF depended on the number of iterations given by the value of m and it is adjusted by the designer. On the other hand, the computational time of the filters based on *discretized linearization* is determined by the ODE solver and the global error control. Considering that the RMSE depicted in Figure 4.2, Figure 4.3 and Figure 4.4 showed that for the 3D-IT-CKF and 5D-IT-CKF there is not an improvement in accuracy after the value of Log2(m) = 4, for that reason there is no need to increase the value for this parameter. Therefore, comparing the values shown in Table 4.2 with Log2(m) = 4 , it can bee seen that there is at least one order of magnitude in speed between the Runga-Kutta and Ito-Taylor implementations.

Lastly, Table 4.2 also confirms that the filters based on the fifth-degree cubature rule are between 3 and 4 times more time consuming when compared with the filters using the third-degree rule.

There are many important conclusions we can get from these results:

- For the filters based on the Ito-Taylor expansion, once the numer of iterations (m) has reached the value of $log_2(m) = 4$, the filters performance does not improve beyond this value; however computation time does increase.

- The filters based on the *discretized linearization* (3D-RK-CKF and 5D-RK-CKF) did not show a better performance and were significantly more time consuming than their *linearized discretization* counterparts (3D-IT-CKF and 5D-IT-CKF). It seems that the Type 2 errors caused by the assumption of the permanent Gaussian over all the time intervals, are bigger than the Type 1 errors made by the discretization of the Ito-Taylor expansion.

- There is an improvement using the 5D-IT-CKF as compared to the 3D-IT-CKF; however, this improvement is small. With that in mind, before using the 5D-IT-CKF, it is important to consider the fourfold increase in computational time compared to the gain in accuracy. There are some applications that will require the maximum accuracy possible, others may not.

- The filters based on the fifth-degree cubature rule showed a bigger number of divergences compared to the filters based on the third-degree cubature rule for large sampling times. This is caused by the lack of an square-root

implementation. With that in mind, the sampling time of the application should be considered before using the fifth-degree filters.

- Whether we use the *discretized linearization* or *linearized discretization* approach, the continuous-discrete cubature Kalman filter use Gaussian approximations that cause errors that cannot be controlled. It would be an interesting topic of research to measure these errors and compare them against filters that do not assume this Gaussianity.

In summary, in this chapter I introduced the *linearized discretization* and the *discretized linearization* approaches of the CD-5D-CKF. I also tested these filters in a challenging scenario with nonlinearities in the process and measurement equations. The results give evidence of a small improvement in accuracy achieved using the fifth-degree cubature rule. In addition, these results revealed the superiority of the *linearized discretization* approach regarding accuracy and computational time, which is a current topic of research. We now move into the second part of the thesis that is related to measurement origin uncertainty.

Chapter 5

Measurement Origin Uncertainty

5.1 Introduction

Measurement origin uncertainty is a phenomena that takes place in remote sensing when the sensor receives measurements whose origin is uncertain (i.e. when it is unknown if the measurements were generated by the target of interest or by the natural environment). When the measurements are originated by the natural environment we refer to them as clutter. Sources of clutter include:

- Backscatter from unwanted objects
  - Sea, rain, ground, ground traffic, birds
- Multipath reflections
- Interference from other radars

Target tracking in clutter is difficult because using spurious measurements in the filter leads to increase in the estimation error (e.g. RMSE in position) and consequently a track loss. Some of the proposed solutions for target tracking in clutter are the strongest-neighbor filter (SNF); Li (1998) and the nearest-neighbor filter (NNF); Li and Bar-Shalom (1996). While the SNF selects the highest intensity measurement from the possible measurements originated by the target, the NNF selects the one that is closest to the predicted measurement. The often used approach has been the probabilistic data association (PDA) technique (Bar-Shalom et al., 2009), which uses a set of validated measurements to estimate the state and covariance of the object rather than selecting and relying upon a single measurement. Another more powerful algorithm is the multiple hypothesis tracker (MHT), which associates all the measurements to the tracks; gating and prunning are used to eliminate tracks with a low probability (Blackman, 2004). The MHT has been shown to be very accurate; nevertheless, due to its long computation time it is not often used.

The PDA has been widely chosen as a good balance between accuracy and computational complexity, and due to its relevance it is described in section 5.4. Before addressing the properties of the PDA and the modelling of clutter, it is important to emphasize the concept of robustness in this context; consequently, the section 5.2 is dedicated to this topic.

5.2 Robustness

From a practical perspective, clutter is unavoidable and unpredictable, and its severity depends on multiple factors as mentioned in section 5.1. Knowing this,

robustness is an relevant characteristic for a filter. The definition of robustness made in Haykin (2008) establishes the following:

> *For an adaptive filter to be robust, small disturbances (i.e. disturbances with small energy) can only result in small estimation errors. The disturbances may arise from a variety of sources: internal or external to the filter.*

Clutter is considered one of the disturbances external to the filter and typically, as a consequence, the track loss increases significantly. For that reason, robustness may be the factor of choice. The design of robust systems sometimes requires sacrifices, especially when we refer to optimality. Consequently, we can say the following: if robustness is needed, optimality is not always the most important factor. This has been stated by Hassibi et al. (1999) as:

> *The process of adaptive filtering is much more related to robustness with respect to statistical variations than it is to optimality with respect to a specified statistical model.*

This statement suggests that a sub-optimal method can be utilized with the objective of guarantee robustness in the presence of disturbances. We now move into the modelling of clutter.

5.3 Clutter Model

In radar applications, the surveillance region of the radar consists of N_c resolution cells. A resolution cell is defined as the maximum detection region in which it is impossible to distinguish (resolve) two independent targets, and it is

given by the multiplication of *resolution range* $\times$ *resolution azimuth* (Skolnik, 2008). When the signal in this cell exceeds a certain threshold, a detection is acknowledged. If, due to clutter, a detection is declared, this is called a false alarm.

With the assumption that the detections caused by clutter in each cell are independent, and the probability of such a detection is the probability of a false alarm in each cell is $P_{FA} = p$, then the probability mass function of the number of false alarm in these N_c cells is given by the binomial distribution (Bar-Shalom and Li, 1995). This is explained by the fact that the requirements for a binomial distribution are being satisfied, that is, the N_c independent trials given by the number of cells and the fact that there are only two possible outcomes: false alarm or not false alarm with the respective probabilities (p) and $(1-p)$. Hence, the probability mass function of the number of false alarms is given by the equation

$$p\{n_{FA} = m\} = \begin{pmatrix} N_c \\ \\ \mathrm{m} \end{pmatrix} p^m (1-p)^{N_c - m}. \tag{5.1}$$

The spatial density λ of the surveillance region (V), is defined by the equation

$$\lambda = \frac{E[n_{FA}]}{V} = \frac{N_c\, p}{V}. \tag{5.2}$$

If N_c is large and $p \ll 1$, equation (5.1) can be approximated by a Poisson distribution, in the form

$$p\{n_{FA} = m\} = e^{-N_c p} \frac{(N_c\, p)^m}{m!}. \tag{5.3}$$

Using equations (5.2) and (5.3), we can describe the number of false alarms in terms of the spatial density:

$$p\{n_{FA} = m\} = e^{-\lambda V}\frac{(\lambda V)^m}{m!}. \tag{5.4}$$

The spatial distribution of the false alarms is considered to be uniform in the radar surveillance region.

5.4 Probabilistic Data Association

The PDA algorithm (Bar-Shalom et al., 2009) calculates the association probabilities of all the measurements that could have been originated by the target of interest. For this reason we can say that the PDA is an all-neighbor data association method. In PDA the following assumptions are taken into consideration:

- Only one target of interest is present.
- Past information about the target through time $k-1$ is in the form of a normal distribution.

$$p[x_{k-1}|Z^{k-1}] = \mathcal{N}[x_{k-1}; \hat{x}_{k-1|k-1}, P_{k-1|k-1}]. \tag{5.5}$$

- The track has been initialized.
- A measurement validation region is established around the predicted measurement in order to select the set of possible measurements for association to the target of interest.

- Among the possible validated measurements, at most one of them can originate from the target (i.e. only one is correct).
- The remaining measurements are assumed to be false alarms caused by various types of clutter (see page 66).
- Target detection occurs independently over time with a known probability P_D.

Due to non-linearity in our system, we decided to use the frame of CD-CKF for the PDA algorithm; having said that, we present the combined algorithm in the next section.

5.4.1 Prediction

Factorize the posterior covariance $P^j_{k-1|k-1}$ as

$$P^j_{k-1|k-1} = S^j_{k-1|k-1} S^j_{k-1|k-1}. \tag{5.6}$$

Next, evaluate the cubature points ($i = 1, ..., 2n$) as follows:

$$X^j_{i,k-1|k-1} = S^j_{k-1|k-1} \xi_i + \hat{x}^j_{k-1|k-1}, \tag{5.7}$$

where the ξ_i are intersections of the n dimensional sphere and coordinate axis given by

$$\begin{aligned} \xi_i &= \sqrt{n} e_i \quad \forall i = 1, ..., n \\ &= -\sqrt{n} e_i \quad \forall i = n+1, ..., 2n. \end{aligned} \tag{5.8}$$

To evaluate the propagated cubature points ($i = 1, 2, ..., 2n$), we use

$$X^{*(j+1)}_{i,k|k-1} = f_d(X^{j}_{i,k-1|k-1}, kT + j\delta), \tag{5.9}$$

where f_d is the noise-free function given by the Ito-Taylor expansion of order 1.5. We can estimate the predicted state as the average of the sums of the propagated cubature points, as shown by

$$\hat{x}^{j+1}_{k-1|k-1} = \frac{1}{2n}\sum_{i=1}^{2n} X^{*(j+1)}_{i,k-1|k-1}. \tag{5.10}$$

The predicted error covariance matrix can be estimated by

$$\begin{aligned} P^{j+1}_{k-1|k-1} =& (X^{*(j+1)}_{k-1|k-1})(X^{*(j+1)}_{k-1|k-1})^T + \frac{\delta^2}{2}\Big[Lf(\hat{x}^{j}_{k-1|k-1}, kT + j\delta)Q^T \\ &+ Q(Lf(\hat{x}_{k-1}, kT + j\delta))\Big] + \frac{\delta^3}{3}(Lf(\hat{x}^{j}_{k-1|k-1}, kT + j\delta)) \\ &\times (Lf(\hat{x}^{j}_{k-1|k-1}, kT + j\delta))^T + \delta Q. \end{aligned} \tag{5.11}$$

Once j has reached m, evaluate the cubature points from $i = 1, 2, ..., 2n$ as follows: Factorize the predicted covariance

$$P^{j}_{k-1|k-1} = S_{k|k-1}S^T_{k|k-1}. \tag{5.12}$$

Then obtain the cubature points as shown by

$$X_{i,k|k-1} = S_{k|k-1}\xi_i + \hat{x}_{k|k-1}. \tag{5.13}$$

Compute the propagated cubature points from $i = 1, 2, ..., 2n$ defined by

$$Z_{i,k|k-1} = h(X_{i,k|k-1}, k). \tag{5.14}$$

The predicted measurement is obtained with the average of the predicted cubature points, as shown by

$$\hat{z}_{k|k-1} = \frac{1}{2n}\sum_{i=1}^{2n} Z_{i,k|k-1}. \tag{5.15}$$

Therefore, we can estimate the innovation covariance matrix by computing

$$P_{zz,k|k-1} = Z^*_{k|k-1} Z^{*T}_{k|k-1} + R_k, \tag{5.16}$$

where the weighted-centered matrix

$$Z^*_{k|k-1} = \frac{1}{\sqrt{2n}}[X_{1,k|k-1} - \hat{x}_{k|k-1} X_{2,k|k-1} - \hat{x}_{k|k-1} ... X_{2n,k|k-1} - \hat{x}_{k|k-1}]. \tag{5.17}$$

5.4.2 Measurement Validation

The measurement validation is a procedure in which the measurements in the radar are tested to select those that have a high probability of pertaining to the desired target. This process is calculated as:

$$\mathcal{V}_k(\gamma) = \Big\{ z : [z - \hat{z}_{k|k-1}]^T (P_{zz,k|k-1})^{-1} [z - \hat{z}_{k|k-1}] \leq \gamma \Big\}, \tag{5.18}$$

where γ is the gate threshold corresponding to the gate probability P_G, that is the probability of the true measurement being inside the validation gate if it is detected (Bar-Shalom and Li, 1995).

The set of validated measurements at time k is denoted as

$$Z_k = \{z_{i,k}\}_{i=1}^{m_k},$$

where $z_{i,k}$ represents the ith validated measurement and m_k is the number of validated measurements.

5.4.3 Data Association

The association probability of $z_{i,k}$ being the correct measurement is shown by

$$\beta_{k,i} = \begin{cases} \frac{\mathcal{L}_i}{1-P_D P_G + \sum_{j=1}^{m_k} \mathcal{L}_j}, & i = 1, ..., m_k, \\ \\ \frac{1-P_D P_G}{1-P_D P_G + \sum_{j=1}^{m_k} \mathcal{L}_j}, & i = 0, \end{cases} \tag{5.19}$$

where $i = 0$ means that none of the measurements are inside the validation region, P_D is the target detection probability, P_G is the gate probability. With the nonparametric PDA the likelihood is obtained by

$$\mathcal{L}_i \triangleq \frac{\mathcal{N}[z_{k,i}; \hat{z}_{k|k-1}, P_{zz}] P_D V_k}{m_k}, \tag{5.20}$$

where V_k is the volume of the validation region.

5.4.4 State Estimation

The state updated equation of the PDA is defined as

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + W_k v_k, \tag{5.21}$$

where W_k is the continuous-discrete Kalman gain, and v_k is the combined innovation given by

$$W_k = \frac{T_{21}}{T_{11}}, \tag{5.22}$$

and

$$v_k = \sum_{i=1}^{m_k} \beta_{k,i} v_{i,k}, \tag{5.23}$$

where $v_{i,k}$ are the innovation terms of each measurement inside the validation region, and $\beta_{k,i}$ is defined in equation (5.19). The covariance associated with the updated state is defined by

$$P_{k|k} = \beta_{k,0} S_{k|k-1} S^T_{k|k-1} + [1 - \beta_{0,k}] P^c_{k|k} + \tilde{P}_k. \tag{5.24}$$

The covariance of the state updated with the correct measurement in equation (5.24) is given by

$$P^c_{k|k} = S_{k|k-1} S^T_{k|k-1} - W_k P_{k,zz} W^T_k, \tag{5.25}$$

and the spread of the innovations term is obtained by

$$\tilde{P}_k \triangleq W_k \Big[\sum_{i=1}^{m_k} \beta_{k,i} v_{k,i} v^T_{k,i} - v_k v^T_k \Big] W^T_k. \tag{5.26}$$

Lastly, we must factorize the associated covariance for the next cycle using

$$P_{k|k} = S_{k|k} S_{k|k}^T. \tag{5.27}$$

To summarize, this chapter focused on the problem of measurement origin uncertainty. First, the problem description was presented. Second, robustness in the context of filtering was defined. Third, the modelling of clutter was described. Lastly, a practical and often used solution was presented. These topics are the bases for the second proposal of the thesis, which is introduced in the Chapter 6.

Chapter 6

Maximum Likelihood Data Association

The PDA is the result of the pure minimum mean-square error (MMSE) approach to tracking and data association (Bar-Shalom et al., 2009). As a result, the PDA uses a weighted average of all possible associations. On the contrary, the maximum likelihood (ML) approach selects one measurement based on the highest likelihood.

On the one hand, the ML method performs a hard decision-selecting one of the validated measurements as being optimal for the state estimate. Although sometimes it is correct, more often than not it produces a wrong track. On the other hand, based on the pure MMSE approach, the PDA makes a weighted average of all the measurements inside the validation gate. Overall it exhibits better performance than the ML method. Nevertheless, for high clutter environments where the number of validated measurements is high, a weighted average

of all the measurements is still not useful because most of the measurements used for the state estimate will be due to clutter. For this reason, we considered that a method that is balanced between these two approaches might perform better than either alone.

6.1 Maximum Likelihood Data Association

The proposed recursive maximum likelihood data association (MLDA) technique is a modified version of the PDA method. The difference is that in order to create some balance between the MMSE and ML approaches, some measurements inside the validation region are ignored and others are selected. This is because as we increase the number of measurements for the state estimate, overall there is less accuracy considering that only one measurement is target originated. At the same time, selecting only one measurement will significantly increase the chance of the state estimate being wrong. Therefore, only a small number of k measurements should be applied for this algorithm. This concept is shown in Figure 6.1.

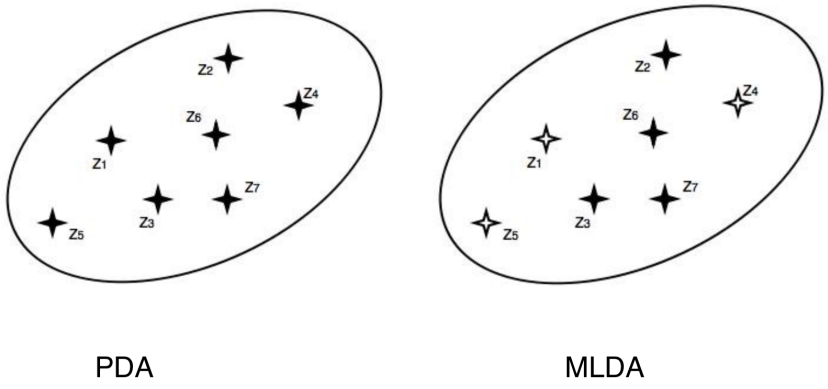

Figure 6.1: On the left, the PDA selects all the measurements (bold crosses) inside the validation region for the updated state. On the right, the MLDA only uses $k = 4$ measurements with the highest or maximum likelihood value for the updated state.

The criterion for choosing which of k measurements to use is based on likelihood value. The best option is to select the ones with the highest values. Fortunately, the PDA technique already calculates the likelihood as shown in equation (5.20). Therefore, all that is needed is to aggregate a sorting algorithm that later selects the values with the maximum likelihood. Once these values have been selected, we can ignore the other measurements and apply the traditional PDA. In short, we are restricting the measurements used by the PDA to the k measurements with the highest likelihood. A block diagram of the proposed MLDA algorithm is shown in Figure 6.2.

With the objective of selecting the best value for the selected k measurements, we performed a black-box optimization. Our aim was to minimize track loss, which is the most common failure while tracking targets in clutter (Korkmaz and Baykal, 2013). With this goal in mind, track loss can be formulated as an unknown function of the selected k measurements and the probability of a

false alarm, which is a measure of the level of clutter. This is described by the equation

$$\text{track loss} = g(k, P_{FA}), \tag{6.1}$$

where g is the black-box. The optimal value for the k measurements is then described as the value that minimizes the track loss, as shown by

$$k_{MLDA} = \text{argmin}\{\text{track loss}\}. \tag{6.2}$$

Therefore, using multiple values for the k measurements under different clutter conditions we can obtain the optimal value for k. The black box optimization is fully described in section 6.1.2.

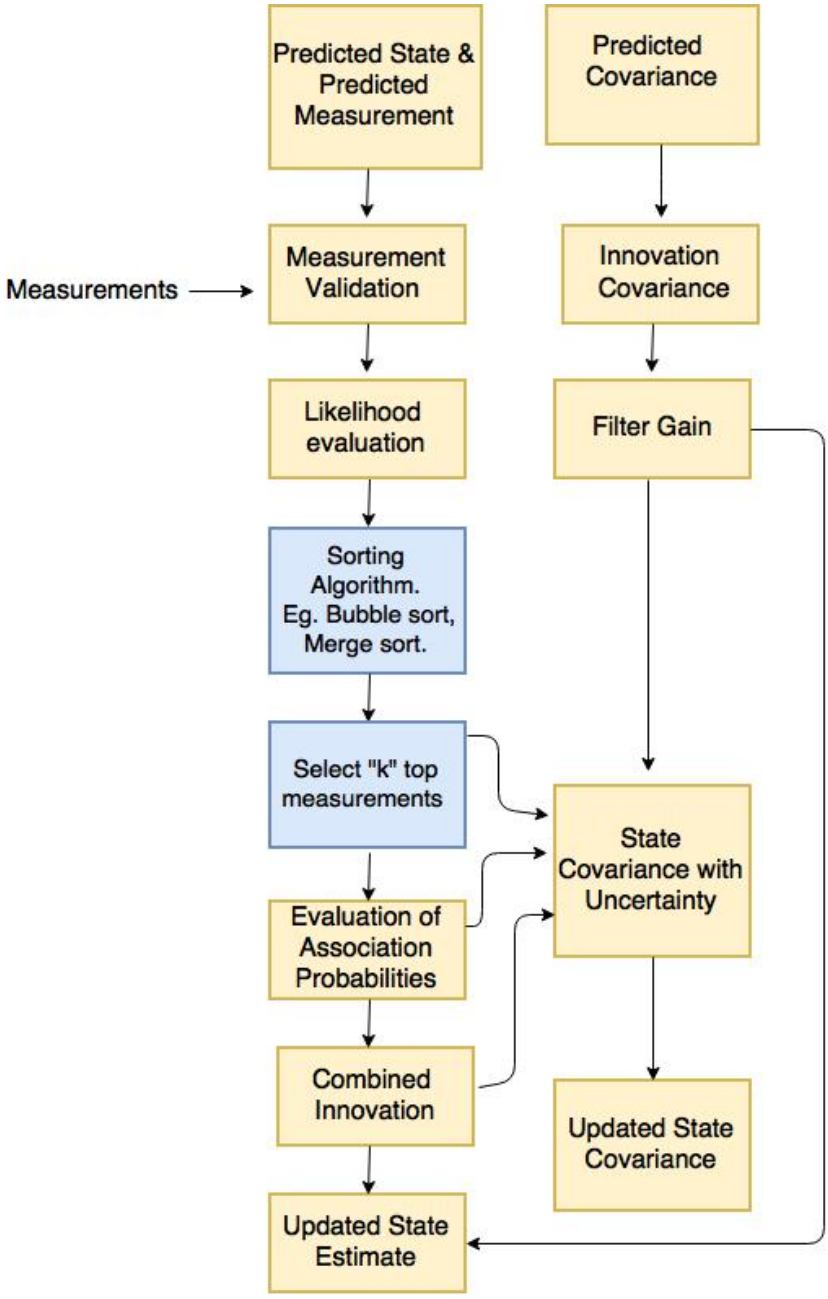

Figure 6.2: One cycle of the Maximum Likelihood Data Association. The blocks in yellow correspond to the traditional PDA and the blocks in blue have been added for the new algorithm.The block diagram of the proposed data association has a similar structure to the one of the PDA; however, two additional blocks are incorporated.

6.2 Black-Box Optimization

As mentioned before, the purpose of the black-box optimization is to find the optimal number of selected measurements k that minimizes the track loss through computer simulations. In light of this, we consider multiple MLDA trackers in which each tracker has a different value of the parameter k. For this optimization, a single target in multiple high-clutter environments is considered.

6.2.1 Target model

The process equation of the target is modelled with a stochastic differential equation given by

$$dx(t) = f(x(t), t)dt + \sqrt{Q}d\beta, \tag{6.3}$$

where $x(t) \coloneqq [x(t) \quad \dot{x}(t) \quad y(t) \quad \dot{y}(t) \quad \omega]^T \in \mathbb{R}^5$ denotes the state of the system at time t with $x(t), y(t)$ and $\dot{x}(t), \dot{y}(t)$ denoting its position and corresponding velocities in Cartesian coordinates. The $\omega(t)$ is the nearly constant turn rate; the drift function $f = [\dot{x}(t), \quad -\omega(t)\dot{y}(t), \quad \dot{y}(t), \quad \omega(t)\dot{x}(t), \quad 0]^T \in \mathbb{R}^5$; β denotes the $n-$dimensional standard Brownian motion with increment $d\beta(t)$ that is independent of $\mathbf{x}(t)$; and Q is the diffusion matrix with $Q = \text{diag}([0, \sigma_1^2, 0, \sigma_1^2, \sigma_2^2])$.

6.2.2 Observation model

The behaviour of the system is observed through noisy measurements sampled at time instant $t_k = kT$ as shown by,

$$z_k = h(x_k) + v_k, \tag{6.4}$$

where $v_k = [w_r, w_\theta]^T$ is the measurement noise with w_r and w_θ being the range and azimuth of the measurement errors respectively, which are assumed to have zero means and the Gaussian known covariance $R = \text{diag}([\sigma_r^2\ \sigma_\theta^2])$; and h is the vector $[r,\ \theta]^T$, with its components described by

$$r = \sqrt{(x_k - x_s)^2 + (y_k - y_s)^2)}, \tag{6.5}$$

$$\theta = \arctan\left\{\frac{y_k - y_s}{x_k - x_s}\right\}, \tag{6.6}$$

where (x_s, y_s) are the Cartesian coordinates of the observer.

The true initial state of the target is $x_0 = [10^3\ \ 0\ \ 2650\ \ 150\ \ 4.5°]$, that is, a 2-dimensional target with initial coordinates $(x = 10^3, y = 2650)$, moving with a speed of $150m/s$ with a turn rate of $4.5°/s$. Following Avitzour (1995), the initial distribution of the target estimate is chosen Gaussian with covariance given by the CD-SCKF after 10 time steps with clutter free zone and the same PD. The simulation scenarios are presented in Figure 6.3. Different clutter densities are considered for the experiment and the parameters of each simulation are summarized in Table 6.1. The true initial state is x_0 and a one point initialization method is used for the experiments.

The track is defined as lost when the true measurement is outside the validation region six times or when the track is more than 300 meters from the true target position the same amount of times. Both conditions are applied because it was observed that the validation region tends to increase in size under high clutter density. As a result, the track would be wrong multiple times, even though the validation region was large enough to include the true measurement. I chose

six times as a threshold considering that due to the sampling interval is $T = 6s$, and not receiving any information of the target after 36 s is enough time to declare a track as lost.

Table 6.1: Parameters used in target tracking simulation with a coordinate turn model

Total Time Simulation	360 seconds
Total Number of Cells	54000
Volume of Surveillance Region	$36 \times 10^6 m^2$
Sampling Interval	6s
Observer Position	1500,100
Probability of Detection	0.9
Probability of False Alarm	0.064, 0.072, 0.080
Gate Threshold	16
Difussion Matrix σ_1	$\sqrt{0.2}$
Difusion Matrix σ_2	7×10^{-3}
Range σ_r	50 m
Azimuth σ_θ	0.1°
Turn Rate ω	4.5°

The true trajectory of the target is simulated applying the Itô-Taylor expansion of order 1.5 to the stochastic differential equation (6.3), resulting in the following stochastic difference equation

$$x_k^j = f_d(x_k^j) + \sqrt{Q}w + (Lf(x_k^j))y, \tag{6.7}$$

with a time step of sampling interval/1000, where the functions Lf and f_d are the following:

$$f_d = \begin{bmatrix} x + \delta\dot{x} - \frac{\delta^2}{2}\omega\dot{(y)} \\ \dot{x} - \delta\omega\dot{y} - \frac{\delta^2}{2}\omega^2\dot{x} \\ y + \delta\dot{y} + \frac{\delta^2}{2}\omega\dot{x} \\ \dot{y} + \delta\omega\dot{x} - \frac{\delta^2}{2}\omega^2\dot{y} \\ \omega \end{bmatrix} \qquad Lf = \begin{bmatrix} 0 & \sigma_1 & 0 & 0 & 0 \\ 0 & 0 & -\sigma_1\omega & 0 & \sigma_2\dot{y} \\ 0 & 0 & 0 & \sigma_1 & 0 \\ 0 & \sigma_1\omega & 0 & 0 & -\sigma_2\dot{y} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

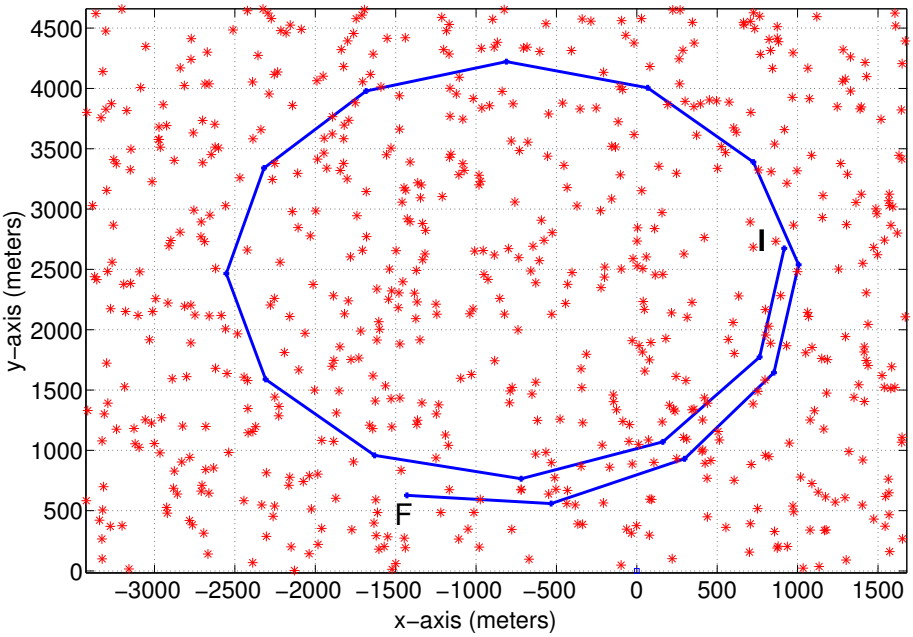

Figure 6.3: Target trajectory in blue. Sea clutter is shown as red with $P_{FA} = 0.044$. There were 16 scans and the sampling was T=6 s. The initial point of the target is denoted with "I" and the final position is denoted with "F".

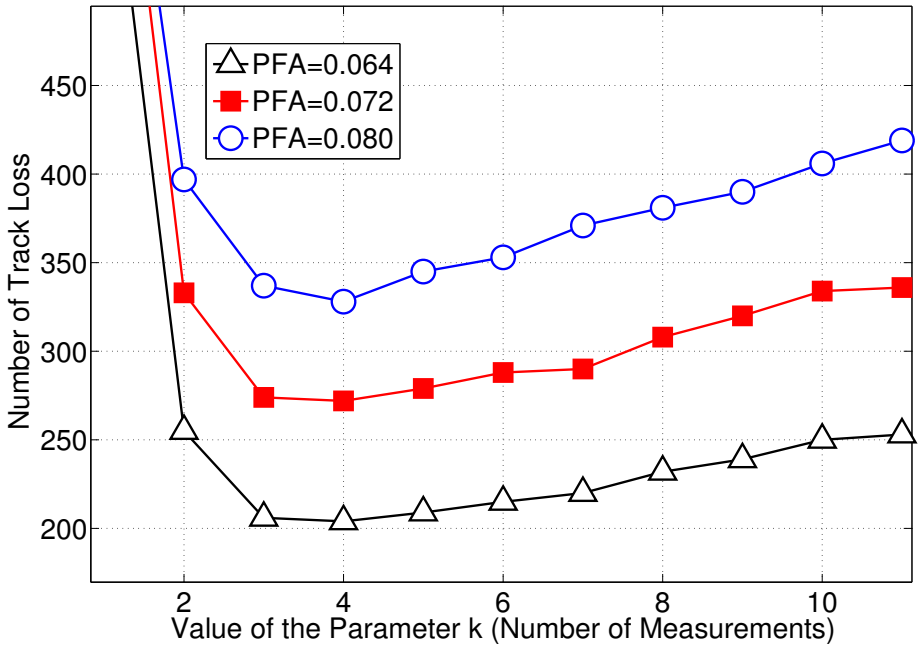

Figure 6.4: Number of track losses in 1000 Monte Carlo simulations for the MLDA with $k = 1, 2, .., 12$ Three different clutter densities were used in the experiment and are shown as false alarms probabilities $P_{FA} = 0.064, 0.072$ and 0.080 shown in black, reed and blue, respectively.

Referring to Figure 6.4, the track loss of multiple MLDA estimators running at the same time is presented. Each of the estimators uses different k selected measurements ranging from one to twelve. For these twelve values of k, we applied three different levels of clutter ($P_{FA} = 0.064, 0.072$ and 0.080). The values of k are represented by the x axis and the track loss by the y axis. From Figure 6.4, we observed three important points:

1. The MLDA showed the same qualitative track loss behaviour in the different clutter densities.

2. The minimum track loss was achieved when the value of k equaled four in all of the different clutter densities. Hence, $k = 4$ is the optimal value.

Once $k > 4$, the algorithm starts increasing the number of track losses as expected, and when $k < 4$, the number of track losses also increases.

3. The MLDA continued performing well when $k = 3$, and the number of track losses was very close to the optimal at $k = 4$.

Once the optimal value of k was obtained, we proceeded to compare the MLDA with the traditional PDA.

6.3 Experimental Results

In this section we report the experimental results obtained by using the MLDA when applied to two nonlinear problems. In the first experiment, we used the coordinated turn model scenario that was used for the black-box optimization. In the second scenario, we modeled one target moving in a linear trajectory under different high clutter conditions.

6.3.1 Experiment 1 - Target with Turn Trajectory

For the first experiment, we used the coordinated turn trajectory with the parameters shown in Table 6.1. The trajectory of the object is shown in Figure 6.3. The track loss of both algorithms under multiple levels of clutter is presented in Figure 6.5.

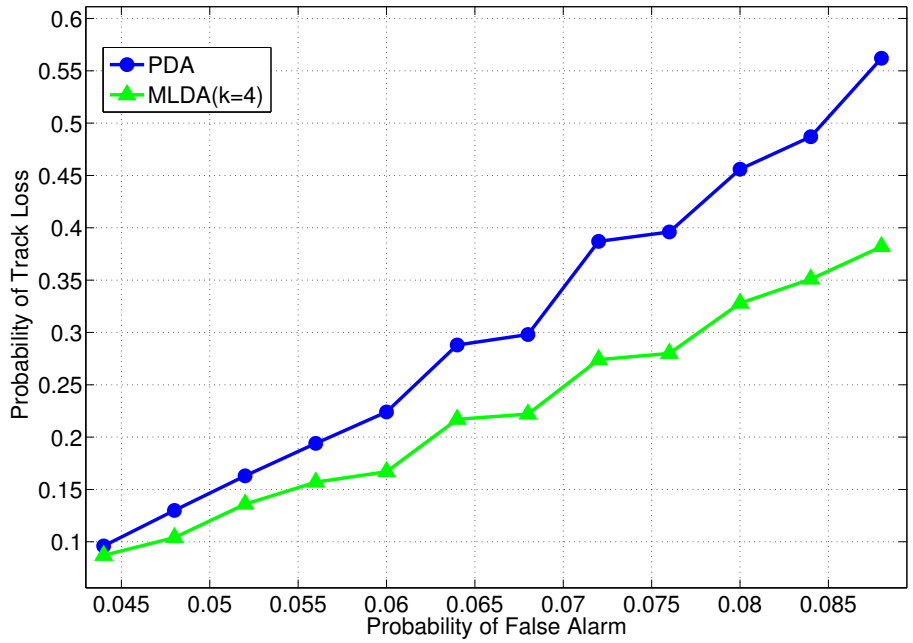

Figure 6.5: Percentage of track loss for the PDA (blue circles) and MLDA (green triangles) with k=4. Results based on 1000 Monte Carlo simulations.

In Figure 6.5, the PDA is compared with the MLDA with $k = 4$ the optimal value from the black-box optimization (Fig. 6.4). The following remarks can be made:

First, at a low probability of false alarm, the number of validated measurements is small. Under these conditions, the MLDA reduces to the standard PDA and the track loss of both algorithms is nearly identical. Second, as the clutter density increases the track loss of both algorithms increases, as expected; however, the robustness of the MLDA is exposed and its track loss is considerably lower than the PDA. This is shown by the diverging slopes of each function, with the PDA slope increasing relative to the MLDA slope as the P_{FA} increases.

6.3.2 Experiment 2 - Target with Linear Trajectory

For the second experiment, I used a single target moving in a linear trajectory under different clutter densities. The target moves with a constant speed and it is sampled by the radar every $T = 2$ s. This simulation scenarios are presented in Figure 6.6. The state of the target is given by $x(t) = [x(t) \quad \dot{x}(t) \quad y(t) \quad \dot{y}(t)]^T \in \mathbb{R}^4$. The drift function $f = [\dot{x}(t), \quad 0, \quad \dot{y}(t), \quad 0]^T \in \mathbb{R}^4$. The process equation of the target is given by equation (6.3), the measurements received are range and azimuth given in equations (6.5) and (6.6) and Gaussian noise with variances σ_r and σ_θ respectively. The true initial state of the target was $x_0 = [100 \;\; 22 \;\; 100 \;\; 5 \;\;]$. The parameters of the simulation are depicted in Table 6.2.

Table 6.2: Parameters used in target tracking simulation with linear movement model

Total Time Simulation	300 seconds
Total Number of Cells	54000
Volume of Surveillance Region	$36 \times 10^6 m^2$
Sampling Interval	2s
Observer Position	0,0
Probability of Detection	0.9
Probability of False Alarm	0.044, 0.048, 0.088
Gate Threshold	9.2
Difussion Matrix σ_1	0.1
Range σ_r	50 m
Azimuth σ_θ	0.1°

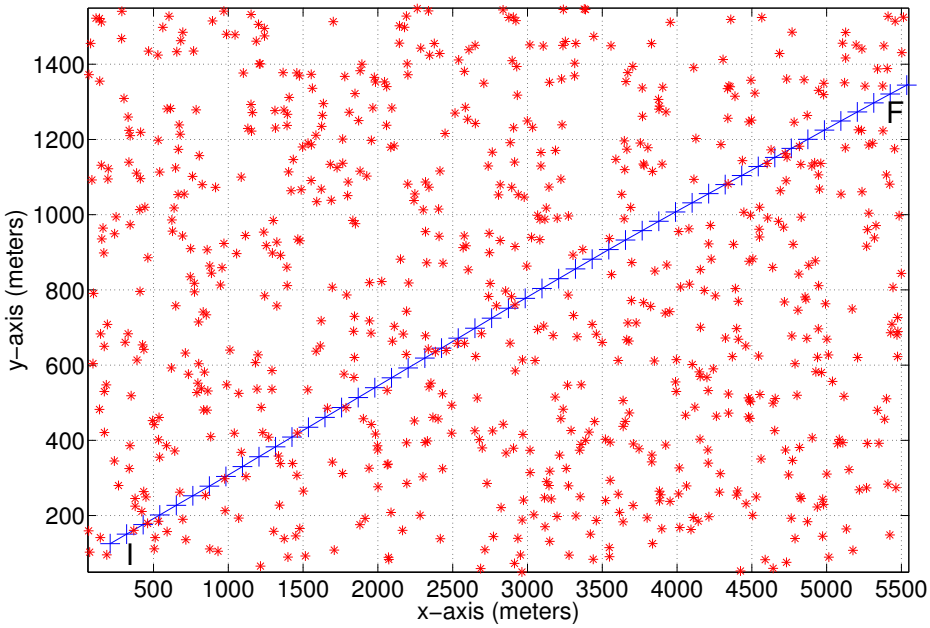

Figure 6.6: Target trajectory in blue. Sea clutter is shown as red with $P_{FA} = 0.048$. There were 60 scans and the sampling was T=2s. The initial position of the target is denoted with "I" and the final position is denoted with "F".

The results shown in Figure 6.7 are consistent with those presented in experiment 1. The MLDA and the PDA have the same track loss at low clutter densities; but as the clutter increases the MLDA technique maintained a lower track loss and outperformed the PDA technique.

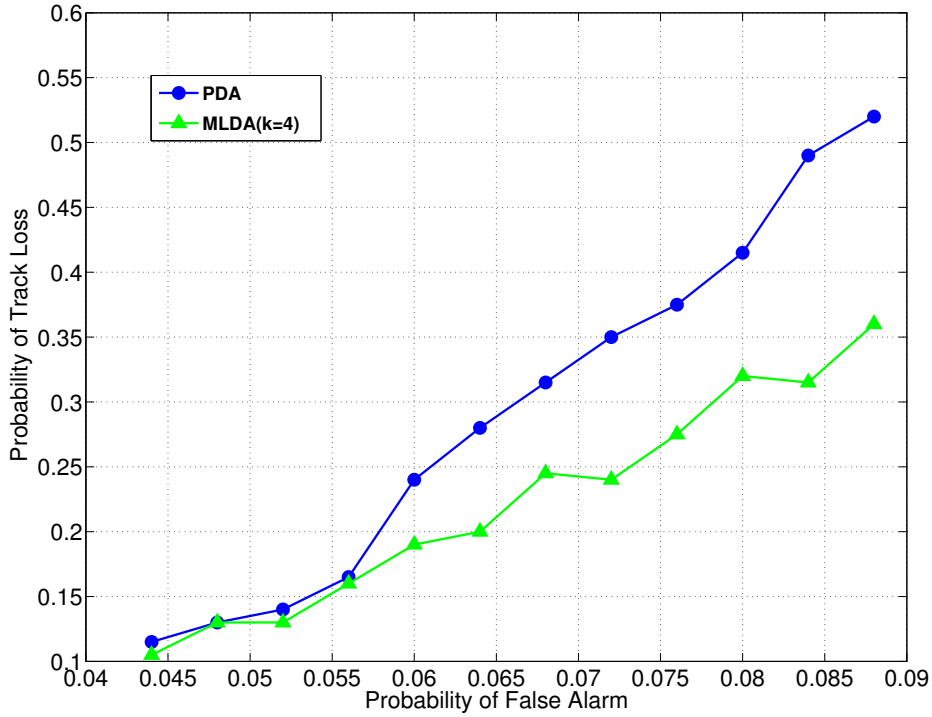

Figure 6.7: Percentage of track losses for the PDA (blue circles) and MLDA (green triangles) with k=4. Results based on 200 Monte Carlo simulations.

6.3.3 Computational Time

For practical applications in tracking, the computational time has always been an issue of concern. This is because target tracking has to operate on a real time; therefore, algorithms should not be computationally too demanding because this delays the calculation and delivery of important results. Considering this requirement, we include in Table 6.3 the time consumption of the MLDA in comparison with the PDA.

Table 6.3 demonstrates that the increment in computational time of the MLDA in comparison with the PDA is approximately 5 percent or less; this is because

the sorting algorithm involved in the MLDA is only done for the validated measurements, which is usually a small number.

Table 6.3: Computational time in seconds of the PDA and MLDA algorithms in experiment 2 as function of two false alarm probabilities for one simulation run using a MacBook Pro with a 2.7 GHz Intel i5 and 16GB DDR3

P_{FA}	PDA	MLDA
0.072	0.7465 s	0.7717 s
0.080	0.8229 s	0.8674 s

6.4 Remarks

In this chapter, I have presented a modified version of the probabilistic data association (PDA) and named the algorithm as the Maximum Likelihood Data Association (MLDA). The proposed algorithm is designed for target tracking under high cluttered environments and has the following desirable characteristics:

- The MLDA is based on the accuracy and stability of the continuous-discrete square-root cubature Kalman filter. This filter has proved to be a good choice for applications with continuous state space models and discrete time measurements as required for target tracking in radar.
- The MLDA is proposed as a mixture of the minimum mean-square error, and maximum likelihood approaches. It is a suboptimal method that has superior robustness compared to the PDA.
- The MLDA uses a small number of measurements for the state estimate, when high clutter is presented. The value of this parameter was obtained

through several experiments with the objective of minimizing track loss. This feature is the key element of the algorithm for improving robustness against high cluttered environments.

- In the case of low clutter densities, the number of validated measurements will typically be less than or equal to the parameter k. Under these conditions, the MLDA reduces to the PDA and good tolerance against clutter is maintained.

- The MLDA algorithm does not require any special adjustments depending on clutter density.

- As was shown in section 6.3.3, the increase in computational time for the MLDA compared to the PDA technique is minimal. This happens because the likelihood values of the measurements have already been calculated by the PDA. The additional operations performed by the MLDA include a sorting algorithm for the validated measurements, which is usually small. Therefore, the sorting can be computed in short time.

These remarks concludes the result chapter and the second proposal of this thesis. The next chapter is dedicated to a brief synopsis of the topics presented and discusses future research perspectives.

Chapter 7

Synopsis and Future Research

7.1 Synopsis

Target tracking has been widely studied over the last decades due to its utility in relevant applications such as navigation and guidance systems, among others. A final and optimal solution has yet to be achieved due to the complex nature of target tracking. The purpose of this research has been to perform the most accurate and robust radar target tracking possible to date. To achieve this goal, this thesis focused on the two main features that are relevant for the accuracy of a radar tracker: the filtering and the data association process. In this regard, I will reiterate the major contributions of my research. In this thesis:

- A new approximate Bayesian filter was proposed for continuous-discrete systems based on the successful cubature Kalman filter and the two known approaches for dealing with the stochastic differential equations presented in the modeling of continuous-time systems.

- I conducted a performance comparison between the *linearized discretization* and *discretized linearization* methods for the continuous-discrete third-degree and fifth-degree cubature Kalman filters.
- Different sources of errors presented on the two known approaches for continuous-discrete filtering were classified.
- I proposed the use of maximum likelihood values for data association to increase the robustness of the target tracking algorithms. Robustness was defined in Chapter 5.
- A black-box optimization method was used to select the smallest number of measurements for data association.
- Challenging computer simulations were conducted and provide evidence of the benefits of the *linearized discretization* approach over the *discretized linearization* method, and that demonstrate the superiority of the current proposals compared to previous state of the art techniques.
- Multiple computer experiments confirmed that the track loss reduction achieved with the proposed data association method (MLDA) was superior compared to the PDA.

Approximate Bayesian filtering and data association are essential elements for radar target tracking; however, the applications of these proposals go far beyond this field and can be used in other areas such as chemical or biomedical engineering. In addition, an important characteristic is the short computational time. In filtering algorithms, it is common to have them to applied in real-time, thus computational demanding methods are often discarded to reduce time requirements. The proposals outlined in this thesis resulted in a small increase

in the computational time compared to traditional techniques used in radar target tracking.

In specific, the proposed 5D-CD-CKF has an increase in computational time compared to the third-degree CD-CKF due to the number of cubature points required. Considering that the reported increase in accuracy is small, it would be important to think carefully if it is necessary to achieve the maximum accuracy possible or use the faster third-degree CD-CKF.

In the case of the proposed MLDA, the increase of the computational time is approximately 5 percent. I believe that this increase in time is acceptable for most of applications to increase the robustness of the state estimation.

7.2 Future Research

7.2.1 Square-root Filtering

Square-root versions of approximate Bayesian filters have been proposed in the literature with the objective to maintain fundamental properties of the covariance matrix, that is, symmetry and positive definiteness (Haykin, 2008). Square-root filters operate with the principle of propagating the square-root of the covariance matrix instead of the covariance matrix itself. The purpose of propagating the square-root matrix is to avoid operations that can result in the loss of the symmetry or positive definiteness such as: matrix square rooting, matrix inversion, and subtraction of two positive-definite matrices. Unfortunately, some of these operations form part of the proposed CD-5D-CKF,

reducing its stability. For this reason, the design of square-root version of the CD-5D-CKF could give more stability to the filter proposed in this thesis.

7.2.2 Multi-Target tracking

Dealing with multiple targets is more challenging than working with a single one. Specifically, when two or more targets are near each other, the association probabilities of the measurements are not independent and joint association probabilities have to be computed (Bar-Shalom et al., 2009). The joint probability data association filter (JPDA) is an extension of the PDA which computes these joint association probabilities when multiple targets are presented.

Since the data association algorithm proposed in this thesis was designed for a single target and the proposed method is a modification to the PDA, the next logical step is to apply the same modification, and thus use maximum likelihood values to the JPDA, and then analyze the performance of the new method in environments with multiple targets.

7.2.3 Self-Driving Cars

Fully autonomous self-driving cars has been set as one of the technical goals for the next generation of automobile. Considering this, researchers in multiple areas of engineering have proposed methods to solve this challenging problem. One of the components in these current proposals is to use radar target tracking to measure the distance between the self-driving car and the surrounding cars with the objective of accelerating or decelerating as needed (Attari et al., 2016). Target tracking on roads has very specific conditions of resolution and clutter

that are different for the ones considered in the air traffic control scenario used in the computer experiments of this thesis. For example, driving while it is snowing or raining can introduce high measurement origin uncertainty. Due to the high risk presented in autonomous driving, I think the proposed data association method might be beneficial for this application due to the high robustness it demonstrated. With this in mind, it would be important to test the proposed data association algorithm in a variety of road conditions.

7.2.4 Cognitive Dynamic Systems

Cognitive dynamic systems (CDS) have been developed in the last 12 years with the objective to design engineering systems that possess characteristics of the human brain. To achieve this goal, CDS have been inspired by advances in understanding the human brain by researchers in neuroscience, as demonstrated by the work of Joaquin Fuster (Fuster, 2003). Given this knowledge, there are at least four principles of cognition observed in Fuster's work that are now considered when designing CDS: perception-action cycle, memory, attention, and intelligence (Haykin, 2012). In any CDS, there are two main components: the perceptual processing algorithm, known as the perceptor, and an algorithm responsible for executing actions, known as the executive. These components are linked through a feedback channel, creating with this, the perception-action cycle. This is shown in Figure 7.1

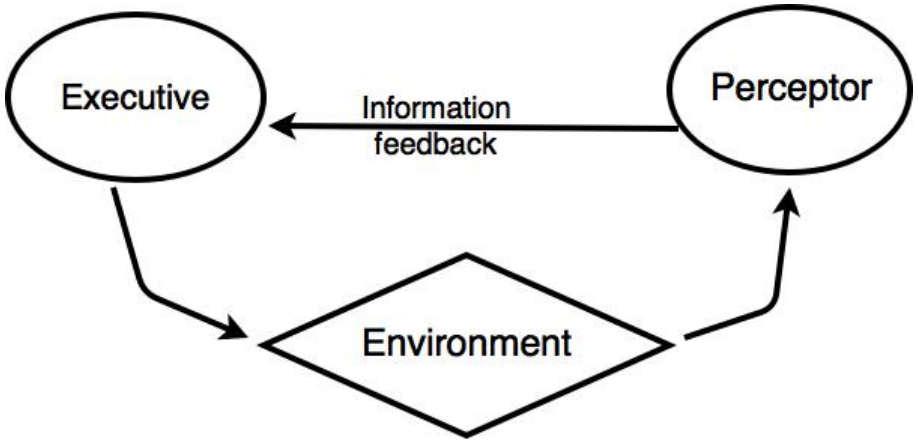

Figure 7.1: Diagram illustrating the concept of the perception-action cycle.

The perceptor, which is responsible for observing and collecting prior information about the world, has to extract information from the measurements or observables and send this information to the executive. The executive has the responsibility of making a decision based on the present information received, prior information, and previous experience (Haykin et al., 2012a). During the development of the CDS, different algorithms have been proposed for the perceptor. For example, in cognitive radar, the perception was in charge of the basic discrete-time CKF (Haykin et al., 2012b), and during the development of cognitive control, the perceptor remained the same (Fatemi and Haykin, 2014). Keeping this in mind, the next generation of CDS could possibly use the continuous-discrete fifth-degree cubature Kalaman filter (CD-5D-CKF) and the new data association method proposed here. It is my sincere desire that these tools will bring robustness and accuracy to future developments in CDS.

Appendix A

Joint Gaussian Random Variables

If two random variables $x \in R^n$ and $y \in R^n$ have Gaussian probability distributions

$$x \backsim N(\bar{x}, P_{xx}), \tag{A.1}$$

and

$$y \backsim N(\bar{y}, P_{yy}), \tag{A.2}$$

then the joint distribution of the vector $z = \begin{bmatrix} x \\ y \end{bmatrix}$ is a normal distribution with the parameters

$$z \backsim N(\bar{z}, P_{zz}), \tag{A.3}$$

with mean and covariance matrices

$$\bar{z} = \begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \tag{A.4}$$

and

$$\bar{P_{zz}} = \begin{bmatrix} P_{xx} & P_{xy} \\ P_{yx} & P_{yy} \end{bmatrix}, \tag{A.5}$$

where

$$P_{xx} = E[(x - \bar{x})(x - \bar{x})^T], \tag{A.6}$$

$$P_{yy} = E[(y - \bar{y})(y - \bar{y})^T], \tag{A.7}$$

$$P_{xy} = E[(x - \bar{x})(y - \bar{y})^T], \tag{A.8}$$

and

$$P_{yx} = E[(y - \bar{y})(x - \bar{x})^T]. \tag{A.9}$$

If y and x are related as

$$y = H_k x + w_k, \tag{A.10}$$

and w_k is a zero-mean Gaussian noise, the terms P_{xy} and P_{yx} are given by

$$P_{xy} = P_{xx} H_k^T, \tag{A.11}$$

and

$$P_{yx} = H_k P_{xx}^T. \tag{A.12}$$

Appendix B

Conditional Gaussian Random Variables

The conditional probability density function (pdf) of x given y when they are jointly Gaussian is given by

$$\begin{aligned} p(x|y) &= \frac{p(x,y)}{p(y)} \\ &= \frac{\sqrt{|2\pi P_{zz}|}\exp(-\frac{1}{2}(z-\bar{z})^T P_{zz}^{-1}(z-\bar{z}))}{\sqrt{|2\pi P_{yy}|}\exp(-\frac{1}{2}(y-\bar{y})^T P_{yy}^{-1}(y-\bar{y}))}. \end{aligned} \tag{B.1}$$

where $z = [x^T \;\; y^T]^T$, using the change of variables

$$\chi = x - \bar{x}, \tag{B.2}$$

and

$$\psi = y - \bar{y}, \tag{B.3}$$

to change the variables x and y to a zero mean variables the exponent on the right side of equation (B.1) has the following form

$$\begin{aligned} q &= \begin{bmatrix} \chi \\ \psi \end{bmatrix}^T \begin{bmatrix} P_{xx} & P_{xy} \\ P_{yx} & P_{yy} \end{bmatrix}^{-1} \begin{bmatrix} \chi \\ \psi \end{bmatrix} - \psi^T P_{yy}^{-1} \psi \\ &= \begin{bmatrix} \chi \\ \psi \end{bmatrix}^T \begin{bmatrix} T_{xx} & T_{xy} \\ T_{yx} & T_{yy} \end{bmatrix}^{-1} \begin{bmatrix} \chi \\ \psi \end{bmatrix} - \psi^T P_{yy}^{-1} \psi \end{aligned}, \tag{B.4}$$

where

$$T_{xx}^{-1} = P_{xx} - P_{xy} P_{yy}^{-1} P_{yx}, \tag{B.5}$$

$$P_{yy}^{-1} = T_{yy} - T_{yx} T_{xx}^{-1} T_{yx}, \tag{B.6}$$

and

$$T_{xx}^{-1} T_{xy} = -P_{xy} P_{yy}^{-1}. \tag{B.7}$$

The exponent q in equation (B.5) can be expanded as

$$\begin{aligned} q &= \chi^T T_{xx} \chi + \chi^T T_{xy} \psi + \psi^T T_{yx} \chi + \psi^T T_{yy} \psi - \psi^T P_{yy}^{-1} \psi \\ &= (\chi + T_{xx}^{-1} T_{xy} \psi)^T T_{xx} (\chi + T_{xx}^{-1} T_{xy} \psi) + \psi^T (T_{yy} - T_{yx} T_{xx}^{-1} T_{xy}) \psi - \psi^T P_{yy}^{-1} \psi \\ &= (\chi + T_{xx}^{-1} T_{xy} \psi)^T T_{xx} (\chi + T_{xx}^{-1} T_{xy} \psi). \end{aligned} \tag{B.8}$$

Substituting equations (B.2), (B.3) and (B.7) into equation (B.8) yields

$$\chi + T_{xx}^{-1} T_{xy} \psi = x - \bar{x} - P_{xy} P_{yy}^{-1} (y - \bar{y}). \tag{B.9}$$

This result shows that the conditional density $p(x|y)$ is Gaussian with the

corresponding mean of

$$\mathrm{E}[x|y] = \bar{x} + P_{xy}P_{yy}^{-1}(y - \bar{y}), \tag{B.10}$$

and the corresponding covariance of

$$\mathrm{cov}[x|y] = T_{xx}^{-1} = P_{xx}^{-1} - P_{xy}P_{yy}^{-1}P_{yx}. \tag{B.11}$$

Bibliography

I. Arasaratnam. *Cubature Kalman Filtering Theory & Applications*. PhD thesis, 2009.

I. Arasaratnam and S. Haykin. Cubature Kalman filters. *IEEE Transactions on Automatic Control*, 54(6):1254–1269, 2009.

I. Arasaratnam, S. Haykin, and T. R. Hurd. Cubature Kalman filtering for continuous-discrete systems: theory and simulations. *IEEE Transactions on Signal Processing*, 58(10):4977–4993, 2010.

M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2002.

K. J. Åström. *Introduction to stochastic control theory*. Courier Corporation, 2012.

M. Attari, Z. Luo, and S. Habibi. An svsf-based generalized robust strategy for target tracking in clutter. *IEEE Transactions on Intelligent Transportation Systems*, 17(5):1381–1392, 2016.

D. Avitzour. Stochastic simulation bayesian approach to multitarget tracking. *IEE Proceedings-Radar, Sonar and Navigation*, 142(2):41–44, 1995.

Y. Bar-Shalom and X.-R. Li. *Multitarget-multisensor tracking: principles and techniques*, volume 19. London, UK:, 1995.

Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.

Y. Bar-Shalom, F. Daum, and J. Huang. The probabilistic data association filter. *IEEE Control Systems*, 29(6), 2009.

V. Beneš. Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics: An International Journal of Probability and Stochastic Processes*, 5(1-2):65–92, 1981.

S. S. Blackman. Multiple hypothesis tracking for multiple target tracking. *IEEE Aerospace and Electronic Systems Magazine*, 19(1):5–18, 2004.

D. Crouse. Basic tracking using nonlinear continuous-time dynamic models [tutorial]. *IEEE Aerospace and Electronic Systems Magazine*, 30(2):4–41, 2015.

J. R. Dormand and P. J. Prince. A family of embedded runge-kutta formulae. *Journal of Computational and Applied Mathematics*, 6(1):19–26, 1980.

A. Doucet, N. De Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice. Series Statistics For Engineering and Information Science*. Springer New York, 2001.

E. B. Dynkin. Markov processes. In *Markov Processes*, pages 77–104. Springer, 1965.

B. Everitt and A. Skrondal. *The Cambridge dictionary of statistics*, volume 106. Cambridge University Press Cambridge, 2002.

M. Fatemi and S. Haykin. Cognitive control: Theory and application. *IEEE Access*, 2:698–710, 2014.

A. Fuller. Analysis of nonlinear stochastic systems by means of the fokker–planck equation. *International Journal of Control*, 9(6):603–655, 1969.

J. M. Fuster. *Cortex and mind: Unifying cognition.* Oxford University Press, 2003.

A. Genz. Fully symmetric interpolatory rules for multiple integrals over hyperspherical surfaces. *Journal of Computational and Applied Mathematics*, 157 (1):187–195, 2003.

N. J. Gordon, D. J. Salmond, and A. F. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *IEE Proceedings F (Radar and Signal Processing)*, 140(2):107–113, 1993.

M. S. Grewal and A. P. Andrews. Applications of kalman filtering in aerospace 1960 to the present [historical perspectives]. *IEEE Control Systems*, 30(3): 69–78, 2010.

B. Hassibi, A. H. Sayed, and T. Kailath. *Indefinite-Quadratic Estimation and Control: A Unified Approach to H2 and H-infinity Theories*, volume 16. SIAM, 1999.

S. Haykin. *Cognitive dynamic systems: perception-action cycle, radar and radio.* Cambridge University Press, 2012.

S. Haykin, M. Fatemi, P. Setoodeh, and Y. Xue. Cognitive control. *Proceedings of the IEEE*, 100(12):3156–3169, 2012a.

S. Haykin, Y. Xue, and P. Setoodeh. Cognitive radar: Step toward bridging the gap between neuroscience and engineering. *Proceedings of the IEEE*, 100 (11):3102–3130, 2012b.

S. S. Haykin. *Adaptive filter theory*. Pearson Education India, 2008.

Y. Ho and R. Lee. A bayesian approach to problems in stochastic estimation and control. *IEEE Transactions on Automatic Control*, 9(4):333–339, 1964.

A. H. Jazwinski. *Stochastic processes and filtering theory*. New York: Academic Press, 1970.

B. Jia, M. Xin, and Y. Cheng. High-degree cubature Kalman filter. *Automatica*, 49(2):510–518, 2013.

S. Julier, J. Uhlmann, and H. F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482, 2000.

R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Journal of Basic Engineering*, 83(1):95–108, 1961.

R. E. Kalman et al. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

A. N. Kolmogorov. Sur l'interpolation et extrapolation des suites stationnaires. *CR Acad. Sci*, 208:2043–2045, 1939.

Y. Korkmaz and B. Baykal. Track loss versus computation time dilemma in multitarget ground target tracking performance. In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 2168–2176. IEEE, 2013.

G. Y. Kulikov. Cheap global error estimation in some runge–kutta pairs. *IMA Journal of Numerical Analysis*, 33(1):136–163, 2013.

G. Y. Kulikov and M. V. Kulikova. The accurate continuous-discrete extended kalman filter for radar tracking. *IEEE Transactions on Signal Processing*, 64(4):948–958, 2016.

X. R. Li. Tracking in clutter with strongest neighbor measurements. i. theoretical analysis. *IEEE Transactions on Automatic Control*, 43(11): 1560–1578, 1998.

X. R. Li and Y. Bar-Shalom. Tracking in clutter with nearest neighbor filters: analysis and performance. *IEEE Transactions on Aerospace and Electronic Systems*, 32(3):995–1010, 1996.

X. Lin, Y. Bar-Shalom, and T. Kirubarajan. Multisensor multitarget bias estimation for general asynchronous sensors. *IEEE Transactions on Aerospace and Electronic Systems*, 41(3):899–921, 2005.

S. Sarkka. On unscented kalman filtering for state estimation of continuous-time nonlinear systems. *IEEE Transactions on Automatic Control*, 52(9): 1631–1641, 2007.

S. Särkkä. *Bayesian filtering and smoothing*. Cambridge University Press, 2013.

S. Särkkä and A. Solin. On continuous-discrete cubature kalman filtering. *IFAC Proceedings Volumes*, 45(16):1221–1226, 2012.

L. F. Shampine and M. W. Reichelt. The matlab ode suite. *SIAM Journal on Scientific Computing*, 18(1):1–22, 1997.

M. Skolnik. *Radar Handbook, Vol. 1, 8.* McGraw Hill Press, 2008.

A. H. Stroud and D. Secrest. *Gaussian quadrature formulas.* Prentice Hall, 1966.

F. Sun and L.-J. Tang. Estimation precision comparison of cubature kalman filter and unscented kalman filter. *Control and Decision*, 28(2):303–308, 2013.

N. Wiener. *Extrapolation, interpolation, and smoothing of stationary time series*, volume 7. MIT Press Cambridge, MA, 1949.