UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

NONLINEAR BAYESIAN ESTIMATION VIA SOLUTION OF THE FOKKER-PLANCK EQUATION

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

JANGHO YOON

Norman, Oklahoma

2009

NONLINEAR BAYESIAN ESTIMATION VIA SOLUTION OF THE FOKKER-PLANCK EQUATION

A DISSERTATION APPROVED FOR THE
SCHOOL OF AEROSPACE AND MECHANICAL ENGINEERING

BY

Dr. Prakash Vedula, Co-Chair

Dr. Yunjun Xu, Co-Chair

Dr. David P. Miller

Dr. Luther White

Dr. Thordur Runolfsson

To my late mother who closed the last chapter of her life two years ago
To my first son who opened the first chapter of his life two years ago
Regretting that they did not have chance to know each other

Acknowledgements

I would like to express my deepest thanks to my dissertation advisors, professors Yunjun Xu and Prakash Vedula. I thank them for their continuous motivation, encouragement, and guidance without which I could not get my dissertation done. I would also like to thank professors David P. Miller, Luther White, and Thordur Runolfsson for their help and for being on my dissertation committee. Most of all, I'd like to thank my family, especially my wife Eunhee for all the love she carried into my life and my father for his graceful support.

Contents

List of Tables

List of Figures

Abstract

A general approach to optimal nonlinear filtering can be described by a recursive Bayesian approach. The key step in this approach is to determine the probability density function of the state vector conditioned on available measurements. However, an optimal solution to the Bayesian filtering problem can only be obtained exactly for a small class of problems such as linear and Gaussian cases. Therefore, in practice, approximate solutions, such as the extended Kalman filter, have been used.

An optimal nonlinear filtering in a recursive Bayesian approach is a two-step process which consists of the prediction and the update process. In the update process, the priori conditional state probability density function (PDF) from the prediction process is updated through Bayes' rule using measurements from sensors. The prediction of conditional state PDF can be made by solving the Fokker-Planck equation (FPE) that governs the time-evolution the conditional state PDF. However, it is extremely difficult to obtain an analytical solution of the Fokker-Planck equation with the exception of a few special cases. So far this estimation method has not been employed much in practice because of the high computational cost needed in solving the FPE numerically. In this dissertation, methods to improve the efficiency of the numerical method in solving the FPE are investigated to enhance the efficiency of the nonlinear filtering.

Two finite difference methods, namely i) the explicit forward method and ii) the alternating direction implicit (ADI) method, are used to solve the FPE numerically. Although the explicit forward method is much simpler to implement, the ADI method is preferred for its low computational cost. To reduce the computational cost further, as the first contribution of the dissertation, a moving domain scheme is developed

to reduce the domain of integration required for solving the Fokker-Planck equation numerically. Simulation results show that the accuracy of the estimation is improved as compared with the Extended Kalman Filter, and at the same time the computational cost is significantly lower with the proposed moving grid scheme than the case without it.

Recently a nonlinear filtering algorithm using a direct quadrature method of moments was proposed, where the associated Fokker-Planck equation is solved efficiently via discrete quadrature based on moment constraints. For some problems, however, this approach showed the phenomenon similar to the "degeneracy" in a particle filter, which is the concentration of weight on particular particles. The possible cause of the phenomenon is that only the weights are updated through the modified Bayes' rule. Therefore, in this dissertation, as another contribution, a new hybrid filter is proposed where the measurement update equations in the extended or the unscented Kalman filter are used along with the direct quadrature method of moments to solve the FPE. In this way the "degeneracy" problem can be mitigated.

Then, new proposed filtering methods are applied to several challenging problems such as i) the bearing-only tracking problem, ii) the relative orbit position estimation problem, and iii) the orbit determination problem to demonstrate their advantages. Simulation results indicate that the performance of the proposed filters are better than existing nonlinear filtering methods, such as the Extended Kalman Filter especially with less measurement updates.

CHAPTER 1

Introduction

This chapter starts with a brief discussion of the importance of the state estimates of a dynamical system from noisy measurements. Then, two major approaches in the Bayesian recursive filter, which is the most commonly used optimal nonlinear filtering method, are reviewed, and the nonlinear filtering via the solution of the Fokker-Planck equation is presented. Finally, the contributions and the organization of this dissertation are presented

1.1 Overview

The problem of finding the state of a system, i.e, what a system is doing, from measurements corrupted by noise is called estimation or filtering [38]. Estimating the state of a dynamical system is very important in engineering since it is necessary to know what a system is doing in order to determine if the system is working properly (monitoring), and to adjust the system to work in the desired manner (controlling). For example, knowing the state of an earth observation satellite, which consists of the position and the velocity of the satellite, is crucial to produce an image of the desired part of the earth. This has been the subject of a considerable amount of research ever since the time Gauss formulated the deterministic least-square technique for a simplified orbit determination problem [26]. To date, many different techniques have been developed and used in a wide variety of applications, such as tracking of a target, ranging with sonar, determining the state of a satellite and an airplane, and estimating the volatility of financial systems using stock market data, etc. [2, 38, 48, 49, 60].

In a broad sense, general approaches to optimal nonlinear filtering can be depicted by the Bayesian approach [3, 36, 67]. In the Baysian approach, one attempts to build the posterior (updated) probability distribution function (PDF) of the state based on the set of measurements and the prior (initial) PDF of the state. This is because the PDF of the state contains all the statistical information about the state. Using this approach, a filter in recursive form, which means that the received measurements are processed sequentially, can be constructed. The Bayesian recursive filtering problem, however, can be solved exactly only for a small class of problems due to the fact that it requires infinite dimensional processes [71]. There are two approaches to nonlinear Baysian filtering problem, which are the local approach and the global approach. In the local approach the filtering problem is approximated so that it can be solved exactly while the global approach attempts to solve the exact problem approximately.

This dissertation deals with the global approach of nonlinear filtering problem based on direct numerical approximations of the optimal nonlinear filter. This can be accomplished through the solution of the Fokker-Planck equation (FPE) in continuous-discrete filtering problems, in which the dynamics are continuous and measurements discrete [38]. Solving the FPE efficiently and accurately is the key to this approach and the focus of the dissertation.

In next section, the details of the local and the global approach are discussed including their advantage and disadvantage followed by the discussion of the nonlinear filtering based on the FPE and the methods to solve the FPE.

1.2 Two Methodologies for Nonlinear Filtering Problems

Methodologies to solve nonlinear Baysian filtering problems can be categorized as i) local approach and ii) global approach [5]. In the local approach, the prior (initial) and the posterior (updated) PDF are approximated with a particular form, usually Gaussian, while in the global approach the posterior density function is calculated

without any explicit assumption about the distribution. So, the nonlinear filters based on the local approach are usually numerically simple, but erroneous when the assumption is violated while the nonlinear filters based on the global approach are numerically complex and expensive, but more accurate than the filters based on the local approach. The prime example of the local approach is the Extended Kalman filter (EKF). The nonlinear system and/or measurement equations associated with the EKF are linearized around the most recent estimation, and PDFs are assumed to be Gaussian. As the EKF utilizes the linearized nonlinear dynamics and/or measurement model achieved through Taylor series expansion around the latest estimation, its performance heavily relies on the degree of nonlinearity [27]. Another example is the Unscented Kalman Filter (UKF). The UKF is built on the nonlinear transformation called the unscented transformation (UT). The UT approximates the mean and covariance of a probability distribution efficiently using a set of points called sigma points. However, these filtering methods have shown the following drawbacks

- When the system and/or measurement models are highly nonlinear, the filter can give a poor or even unstable performance as a consequence of the linearization involved in the filtering algorithm.
- The derivation of the Jacobian, which is part of the linearization process, is not a trivial task in many applications.
- Numerical evaluation of the Jacobian can be computationally intensive.

These drawbacks can be resolved by solving the nonlinear filtering problem using the global approach which is more general framework of nonlinear filtering problem.

The global approach approximates the PDF directly. Therefore, the computational complexity of the global approach can be far greater than a local approach, which has been the major limiting factor for the global approach. However, the enhancement of the performance coming from the global approach may outweigh the

additional computational cost. The particle filter based on the sequential Monte Carlo (SMC) methods [20, 31, 58, 66] is a good example of the global approach. The sequential Monte Carlo filter can be described as a recursive filter based on Monte Carlo simulation schemes with an importance sampling. This technique estimates a distribution from samples generated from a different distribution other than the distribution of interest in order to solve online estimation and prediction problems [22]. The sequential Monte Carlo approach is known as bootstrap filtering [30], condensation algorithm [56], and particle filtering [14]. The flexible nature of the Monte Carlo simulations makes these methods more adaptive [58]. There have been many modifications and improvements recently on particle filters [21]. However, there still exist many problems related to the choice of a proposal distribution, a sampling mechanism from the distribution, and high computational complexity.

1.3 Nonlinear Filtering with the Fokker-Planck Equation

Global approaches for nonlinear filtering are often based on direct numerical approximations of the optimal nonlinear filter, which can be accomplished through the use of numerical methods to solve Kushner equation or Zakai equation in the case of continuous-continuous system models [12, 28, 37, 55], or the Fokker-Planck equation in continuous-discrete system models for the state conditional PDF [15, 18, 38]. The optimal estimates in the sense of the minimum mean square error (MMSE) or the maximum likelihood can be constructed from the approximated posterior density.

The Fokker-Planck equation (FPE), which is also known as the Kolmogorov forward equation, was first used by Adriaan Fokker [24] and Max Planck [62] to explain the Brownian motion of particles in fluid. It describes the evolution of the transition probability density of the Markov process produced by the Itô stochastic differential equation (SDE) [38]. The filter based on the solution of the FPE will be able to perform properly if the FPE is solved quickly and accurately. The analytical solution

of the FPE is difficult to obtain with a few exceptions. Therefore, the evaluation of the FPE between measurements usually has to be done numerically. This has prevented the use of this nonlinear filtering algorithm until recent years. The focus of the dissertation is to find an efficient solution to the nonlinear filtering problem by improving the computational efficiency of the numerical solution of the FPE. This has been achieved in two different ways. The first one is through the finite difference methods with moving domain, and the second one is by employing the direct quadrature method of moments.

Solving the FPE was done using two distinct finite difference methods, namely i) the explicit forward method and ii) the alternating direction implicit (ADI) method. Although the explicit forward method is much simpler to implement, the ADI method is preferred for its low computational cost. The computational cost of using any finite difference method alone will still be very high as it is necessary to use a very large computational domain because the PDF is a function defined over an infinite domain. Consequently, in order to reduce the computational load, it is necessary to have a method that would allow us to reduce the size of the domain without any compromise in the solution accuracy [15, 48, 49]. As one of the contributions of this dissertation, an adaptive moving domain will be described (in Chapter 3).

Even with adaptive grids, the computational cost of the finite difference based numerical method can be still too high for the high dimensional system [16, 77, 78]. So, a new approximation approach based on the direct quadrature method of moments (DQMOM) [6], along with Bayes' formula was studied [75] for solving the FPE based nonlinear filtering problems. Originally, the direct quadrature method of moments was developed to solve the population balance [57]. This approach involves a representation of the state conditional PDF in terms of a finite summation of the Dirac delta functions, whose weights and locations (abscissas) are evolved under the moment constraints and modified by Bayes' rule for the measurement update. Using

a small number of scalars (to describe to weights of Dirac delta functions), the method is able to efficiently and accurately model the stochastic processes through a set of algebraic ordinary differential equations (ODEs). The DQMOM approach could lead to a significant reduction in computational cost, compared to finite difference (and other equivalent) methods, especially for high dimensional problems.

Although the DQMOM approach seems promising, the DQMOM based nonlinear filter with modified Bayes' rule as the measurement update showed that the "degeneracy" phenomenon, similar to the one exists in a typical particle filter because in this algorithm only the weight is updated and the abscissas remain the same [76]. As a solution to this problem, a hybrid approach is used. In this approach, the PDE is propagated through DQMOM algorithm to provide the predicted PDF, and the steps borrowed from the EKF or the unscented Kalman Filter (UKF) perform the measurement update to generate the posterior PDF. Using the UKF update equations offer an additional benefit that it does not require the linearization of the measurement equation.

1.4 Contribution

The purpose of this research is to investigate new efficient nonlinear estimation algorithms such that the computational cost in solving the FPE can be substantially reduced. The specific contributions are

- An efficient and simple adaptive moving domain is developed to reduce the computational cost in solving the Fokker-Planck equation via the finite difference approach.
- A form of the ADI method that can be applied to wide range of the nonlinear dynamic problems has been derived.

- Hybrid filtering algorithms are proposed to mitigate the "degeneracy" phenomenon seen in the DQMOM based nonlinear filtering.

- The proposed filtering algorithms are successfully applied to selected problems such as i) the bearing-only tracking problem, ii) the relative orbit position estimation problem, and iii) the orbit determination problem.

1.5 Organization of the Dissertation

In Chapter 2, we provide a overview of the existing linear/nonlinear filtering methods. In this chapter, the most commonly used filtering methods such as the Extended Kalman filter (EKF) along with the Kalman filter, the unscented Kalman filter (UKF), and the particle filter are reviewed. The advantages and drawbacks of each filter are discussed.

In Chapter 3, the nonlinear filter via the numerical solution of the Fokker-Planck equation is discussed. Brief introduction of the Fokker-Planck equation is presented followed by the discussion of the finite difference methods in the derivation of the ADI method and adaptive domain schemes. The direct quadrature method of moments is introduced, and the update methods are discussed.

In chapter 4, the proposed nonlinear filtering algorithms with applications to the spacecraft relative position estimation, the orbit determination, and the bearing-only tracking problem are investigated, and numerical simulation results are presented. The performance by these filters is compared with the Extended Kalman filter and the Unscented Kalman filter.

Chapter 5 presents the summary of the dissertation and directions for future works are discussed in the end.

CHAPTER 2

Review of Estimation Theories

In this chapter, the most commonly used filtering methods started with Baysian recursive filter are reviewed. For linear filtering methods, linear sequential estimation and Kalman filter are reviewed. Especially, the Kalman filter will be derived through the Fokker-Planck equation and the Bayes' rule to show that the Kalman filter is the special case of the filtering method based on the Fokker-Planck equation and the Bayes' rule. The Extended Kalman filter (EKF), the unscented Kalman filter (UKF), and the particle filter are reviewed as nonlinear filtering methods. The advantages and drawbacks of each filter are discussed.

2.1 Bayesian Recursive Filter

In a broad sense, general approaches to optimal nonlinear filtering can be depicted by a recursive Bayesian approach[3, 36, 67]. In the Bayesian approach, the problem is to find the posterior conditional probability density function (PDF) $p(\mathbf{x}_k|\mathbf{Y}_k)$ where $\mathbf{x}_k \triangleq \mathbf{x}(k)$ is the state vector of the system and $\mathbf{Y}_k \triangleq [\mathbf{y}_0, \mathbf{y}_2, \cdots, \mathbf{y}_k]^T$ is the history of observations. This density function will encapsulate all the information about the state vector $\mathbf{x}_k$ which is contained in the measurement $\mathbf{Y}_k$ and the prior distribution of $\mathbf{x}_{k-1}$, which is the distribution before the measurement update. Once $p(\mathbf{x}_k|\mathbf{Y}_k)$ is found, the optimal estimate can be obtained by the conditional expectation of $\mathbf{x}_k$ given $\mathbf{Y}_k$

$$\hat{\mathbf{x}}_k = E[\mathbf{x}_k|\mathbf{Y}_k] = \int \mathbf{x}_k p(\mathbf{x}_k|\mathbf{Y}_k) d\mathbf{x}_k \tag{2.1}$$

The posterior PDF for the state can be found by the Bayesian recursion relations [36, 67] as

$$p(\mathbf{x}_k|\mathbf{Y}_{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{x}_{k-1}|\mathbf{Y}_{k-1})d\mathbf{x}_{k-1} \tag{2.2}$$

and

$$p(\mathbf{x}_k|\mathbf{Y}_k) = C_k p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Y}_{k-1}) \tag{2.3}$$

where the normalizing constant C_k is

$$C_k = \left(\int p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{Y}_{k-1})d\mathbf{x}_k\right)^{-1} \tag{2.4}$$

With further integration of Eq. (2.2) and C_k, Eq. (2.3) becomes

$$p(\mathbf{x}_k|\mathbf{Y}_k) = \frac{p(\mathbf{x}_k|\mathbf{Y}_{k-1})p(\mathbf{y}_k|\mathbf{x}_k)}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})} \tag{2.5}$$

Equation (2.5) is the Bayes' formula showing that the posterior PDF for $\mathbf{x}$ with given measurements $\mathbf{y}$ is directly proportional to the prior value of $\mathbf{x}$ multiplied by the likelihood of the observation. The PDF $p(\mathbf{y}_k|\mathbf{x}_k)$ is defined by the characteristics of the sensor and usually can be assumed to have a Gaussian distribution. Equations (2.2) and (2.3) form a two-step process or multi-stage estimation [36] to obtain the posterior conditional PDF $p(\mathbf{x}_k|\mathbf{Y}_k)$ [38]. The first step, Eq. (2.2) is to predict $p(\mathbf{x}_k|\mathbf{Y}_{k-1})$, and the second step, Eq. (2.3) is to update it using the measurement $\mathbf{y}_k$ followed by integration of Eq. (3.5) to obtain the state estimation.

However, the equations (2.2) and (2.3) for the Bayesian recursive filtering can only be solved exactly for a small class of problems such as the problems with the linear Gaussian process for the system and measurement model. For such cases, the famous Kalman filter is the optimal solution [44] since it gives the optimal solution in the senses of minimum-mean-square-error (MMSE), maximum likelihood (ML), and maximum a posteriori (MAP) [3, 36, 67]. However, most of problems in engineering

are nonlinear and non-Gaussian, and the exact solution to the recursive Bayesian filtering problem for nonlinear systems is intractable due to the fact that it requires infinite dimensional processes [71].

2.2 Linear Filters

2.2.1 Linear Sequential Estimation

The linear least square method developed by Gauss [26] processes a set/batch of measurements simultaneously to generate an estimate. However, in reality the measurements come in as a stream instead of a batch. In this section, it is assumed that the measurements are taken in a sequence so that the estimates are calculated by utilizing all previous measurements and the current data [9]. The method presented in this section can be found in Ref. [17]. Let's consider the following two sets of sequential measurement vectors and their corresponding measurement equations

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x} + \mathbf{v}_1, \qquad \mathbf{y}_1 = [y_{11}, y_{11}, \cdots, y_{1m_1}] \in \mathbb{R}^{m_1\times 1} \tag{2.6}$$

$$\mathbf{y}_2 = \mathbf{H}_2\mathbf{x} + \mathbf{v}_2, \qquad \mathbf{y}_2 = [y_{21}, y_{21}, \cdots, y_{2m_2}] \in \mathbb{R}^{m_2\times 1} \tag{2.7}$$

where $\mathbf{H}_1 \in \mathbb{R}^{m_1\times 1}$ and $\mathbf{H}_2 \in \mathbb{R}^{m_2\times 1}$ are the linear mapping coefficient matrices. The least-squares estimates $\hat{\mathbf{x}}_1$ of the unknown $\mathbf{x}_1$ based on the first measurement subset (Eq. (2.6)) is obtained by using the equation from the weighted least squares.

$$\hat{\mathbf{x}}_1 = \left(\mathbf{H}_1^T\mathbf{W}_1\mathbf{H}_1\right)^{-1}\mathbf{H}_1^T\mathbf{W}_1\mathbf{y}_1 \tag{2.8}$$

where $\mathbf{W}_1$ is an $m_1\times m_1$ symmetric, positive definite weighting matrix associated with measurement $\mathbf{y}_1$. The measurement set $\mathbf{y}_1$ and $\mathbf{y}_2$ can be merged into one equation

and used simultaneously to find the estimation of $\mathbf{x}_2$.

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v} \tag{2.9}$$

where $\mathbf{y} = [\mathbf{y}_1\ \mathbf{y}_2]^T$, $\mathbf{H} = [\mathbf{H}_1\ \mathbf{H}_2]^T$, $\mathbf{v} = [\mathbf{v}_1\ \mathbf{v}_2]^T$, and the merged weight matrix is assume to be a block diagonal matrix of

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 & 0 \\ 0 & \mathbf{W}_2 \end{bmatrix}$$

The optimal least square estimate based on the two measurements is obtained with Eq. (2.8)

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W}\mathbf{y} \tag{2.10}$$

By expanding the block diagonal matrix $\mathbf{W}$, the equation (2.10) becomes

$$\hat{\mathbf{x}}_2 = \left[\mathbf{H}_1^T\mathbf{W}_1\mathbf{H}_1 + \mathbf{H}_2^T\mathbf{W}_2\mathbf{H}_2\right]^{-1}\left(\mathbf{H}_1^T\mathbf{W}_1\mathbf{y}_1 + \mathbf{H}_2^T\mathbf{W}_2\mathbf{y}_2\right) \tag{2.11}$$

However that is not very efficient approach because this procedure has to be repeated with a new set of measurement, i.e., merging more measurements, to come up with the new equation like Eq. (2.11) for $\hat{\mathbf{x}}_3$. This problem can be solved by the sequential least square method. The core of the sequential approach to the least square problem is to make calculations for the new estimation by efficiently using the previous estimations [17]. To achieve that the following variables are defined

$$\mathbf{P}_1 \triangleq \left[\mathbf{H}_1^T\mathbf{W}_1\mathbf{H}_1\right]^{-1} \tag{2.12}$$

$$\mathbf{P}_2 \triangleq \left[\mathbf{H}_1^T\mathbf{W}_1\mathbf{H}_1 + \mathbf{H}_2^T\mathbf{W}_2\mathbf{H}_2\right]^{-1} \tag{2.13}$$

The relationship between $\mathbf{P}_1$ and $\mathbf{P}_2$ can be established as

$$\mathbf{P}_2^{-1} = \mathbf{P}_1^{-1} + \mathbf{H}_2^T \mathbf{W}_2 \mathbf{H}_2 \tag{2.14}$$

Using this relationship with equations (2.8) and (2.11), the optimal estimate $\hat{\mathbf{x}}_2$ using the previous estimate $\hat{\mathbf{x}}_1$ can be written as

$$\hat{\mathbf{x}}_2 = \hat{\mathbf{x}}_1 + \mathbf{K}_2(\mathbf{y}_2 - \mathbf{H}_2\hat{\mathbf{x}}_1) \tag{2.15}$$

where the optimal gain $\mathbf{K}$ is $\mathbf{K} \triangleq \mathbf{P}_2\mathbf{H}_2^T\mathbf{W}_2$.

These two equations can be generalized to form the recursive least-square estimation which uses the k^{th} estimate to calculate the $(k+1)^{th}$ as

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_k + \mathbf{K}_{k+1}(\mathbf{y}_{k+1} - \mathbf{H}_{k+1}\hat{\mathbf{x}}_k) \tag{2.16}$$

where $\mathbf{K}_{k+1} = \mathbf{P}_{k+1}\mathbf{H}_{k+1}^T\mathbf{W}_{k+1}$, and

$$\mathbf{P}_{k+1}^{-1} = \mathbf{P}_k^{-1} + \mathbf{H}_{k+1}^T\mathbf{W}_{k+1}\mathbf{H}_{k+1} \tag{2.17}$$

This update process is known as Kalman update process [45], and the gain $\mathbf{K}_{k+1}$ is called the Kalman gain matrix. Finding $\mathbf{P}_{k+1}$ involves inverting the $n \times n$ matrix given in Eq. (2.17), which is not always simple. This inversion can be removed using the Sherman-Morrison-Woodbury *matrix inversion lemma* [29], and Eq. (2.17) becomes

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \mathbf{P}_k\mathbf{H}_{k+1}^T\left(\mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T + \mathbf{W}_{k+1}^{-1}\right)^{-1}\mathbf{H}_{k+1}\mathbf{P}_k \tag{2.18}$$

Now, the equation for the Kalman gain can be obtained by substituting the above

equation into $\mathbf{K}_{k+1} = \mathbf{P}_{k+1}\mathbf{H}_{k+1}^T\mathbf{W}_{k+1}$, which yields

$$\begin{aligned}\mathbf{K}_{k+1} &= \left[\mathbf{P}_k - \mathbf{P}_k\mathbf{H}_{k+1}^T\left(\mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T + \mathbf{W}_{k+1}^{-1}\right)^{-1}\mathbf{H}_{k+1}\mathbf{P}_k\right]\mathbf{H}_{k+1}^T\mathbf{W}_{k+1} \\ &= \mathbf{P}_k\mathbf{H}_{k+1}^T\left[\mathbf{I} - \left(\mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T + \mathbf{W}_{k+1}^{-1}\right)^{-1}\mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T\right]\mathbf{W}_{k+1}\end{aligned} \tag{2.19}$$

By factoring $\left(\mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T + \mathbf{W}_{k+1}^{-1}\right)^{-1}$, Eq. (2.19) can be rewritten as

$$\begin{aligned}\mathbf{K}_{k+1} = \mathbf{P}_k\mathbf{H}_{k+1}^T&\left(\mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T + \mathbf{W}_{k+1}^{-1}\right)^{-1} \\ &\times\left[\mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T + \mathbf{W}_{k+1}^{-1} - \mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T\right]\mathbf{W}_{k+1}\end{aligned} \tag{2.20}$$

Since the last term is the identity matrix, the Kalman gain can be written as

$$\mathbf{K}_{k+1} = \mathbf{P}_k\mathbf{H}_{k+1}^T\left(\mathbf{H}_{k+1}\mathbf{P}_k\mathbf{H}_{k+1}^T + \mathbf{W}_{k+1}^{-1}\right)^{-1} \tag{2.21}$$

and, the covariance update equation can be rearranged in terms of the new Kalman gain equation as

$$\mathbf{P}_{k+1} = \left[I - \mathbf{K}_{k+1}\mathbf{H}_{k+1}\right]\mathbf{P}_k \tag{2.22}$$

The state estimation equation (Eq (2.16)) and above two update equations (Eq (2.21) and Eq (2.22)) create the covariance recursion-form.

2.2.2 Kalman Filter

The Kalman filter has been the focus of extensive research since the publication of Kalman's famous paper [44]. The Kalman filter is a mathematical algorithm that generates an efficient recursive solution of the least-squares method. The filter can give an estimate of the past, present, and also future states based on noisy measurements. It produces the minimum variance estimate of the state based on statistical information of the dynamic model and the measurement. The Kalman filter can be derived in several different ways. In this section the Kalman filter will be derived based on

the Fokker-Planck equation and the Bayes' rule to show that the Kalman filter is the special case of the filtering method that is the main topic of this dissertation [38].

Consider the linear Itô stochastic differential equation

$$d\mathbf{x}_t = \mathbf{F}(t)\mathbf{x}_t dt + \mathbf{G}(t)d\boldsymbol{\beta}_t, \quad t \geq t_0 \tag{2.23}$$

where $\mathbf{F}(t) \in \mathbb{R}^{n\times n}$ is a linear state function, $\mathbf{x}_t \in \mathbb{R}^{n\times 1}$ is the state vector, $\boldsymbol{\beta}_t$ is an Brownian motion process with $E[\boldsymbol{\beta}_t\boldsymbol{\beta}_t^T] = \mathbf{Q}(t)dt$, and $\mathbf{G}(t) \in \mathbb{R}^{n\times m}$ is a matrix function. The discrete linear measurement $\mathbf{y}_k$ taken at the discrete time instants t_k is

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k, \quad k = 1, 2, ... \tag{2.24}$$

where $\mathbf{H}_k \in \mathbb{R}^{m\times n}$ is the measurement function, $\mathbf{y}_k \in \mathbb{R}^{m\times 1}$ is the measurement vector, and $\mathbf{v}_k$ is a white Gaussian noise with $N(0, R_k)$. For the given linear system (Eq.(2.23)), the FPE that is responsible for propagation of the PDF between measurements becomes

$$\begin{aligned}\frac{\partial p}{\partial t} &= -\sum_{i=1}^{n}\frac{\partial\left[p\mathbf{F}_i x_i\right]}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2\left[p\left(\mathbf{G}\mathbf{Q}\mathbf{G}^T\right)_{ij}\right]}{\partial x_i \partial x_j} \\ &= -p\mathbf{tr}(\mathbf{F}) - \left(p_x^T\right)\mathbf{F}\mathbf{x} + \frac{1}{2}\mathbf{tr}(\mathbf{G}\mathbf{Q}\mathbf{G}^T)p_{xx}\end{aligned} \tag{2.25}$$

where $p_x \triangleq \frac{\partial p}{\partial \mathbf{x}}$, $p_{xx} \triangleq \frac{\partial^2 p}{\partial \mathbf{x}^2}$, and x_i is the i^{th} component of the state vector $\mathbf{x}$. The characteristic function of a random variable x is the Fourier transform of its probability distribution function

$$\varphi(\mathbf{u}, t) = \int_{-\infty}^{\infty} e^{i\mathbf{u}^T\mathbf{x}} p(\mathbf{x}, t)d\mathbf{x} \tag{2.26}$$

By substitute Eq.(2.25) into the time derivative of Eq.(2.26)

$$\begin{aligned}\frac{\partial\varphi}{\partial t} &= \int_{-\infty}^{\infty} e^{i\mathbf{u}^T\mathbf{x}}\frac{\partial p}{\partial t}\,d\mathbf{x} \\ &= -\varphi\mathbf{tr}(\mathbf{F}) - \int e^{i\mathbf{u}^T\mathbf{x}}p_x^T\mathbf{F}\mathbf{x}d\mathbf{x} + \frac{1}{2}\mathbf{tr}(\mathbf{G}\mathbf{Q}\mathbf{G}^T)\int e^{i\mathbf{u}^T\mathbf{x}}p_x^T p_{xx}d\mathbf{x} \\ &= -\varphi\mathbf{tr}(\mathbf{F}) - \mathbf{tr}(\mathbf{F})\int \mathbf{x}p_x^T e^{i\mathbf{u}^T\mathbf{x}}d\mathbf{x} + \frac{1}{2}\mathbf{tr}(\mathbf{G}\mathbf{Q}\mathbf{G}^T)\int p_{xx}e^{i\mathbf{u}^T\mathbf{x}}d\mathbf{x}\end{aligned} \tag{2.27}$$

Assuming that p and its partial derivatives with respect to x vanish at ∞, calculating the integrals from above equation by part would yield

$$\frac{\partial\varphi}{\partial t} = \mathbf{u}^T\mathbf{F}\frac{\partial\varphi}{\partial\mathbf{u}} - \frac{1}{2}\varphi\mathbf{u}^T\mathbf{G}\mathbf{Q}\mathbf{G}^T\mathbf{u} \tag{2.28}$$

Since the system is assumed to be the Gaussian process, its characteristic function is known to be

$$\varphi(\mathbf{u},t) = exp\left(i\mathbf{u}^T\hat{\mathbf{x}}_\mathbf{t} - \frac{1}{2}\mathbf{u}^T\mathbf{P}_t\mathbf{u}\right) \tag{2.29}$$

Taking the derivatives of the characteristic function with respect to time and $\mathbf{u}$ will give

$$\begin{aligned}\frac{\partial\varphi}{\partial t} &= \varphi\left[i\mathbf{u}^T\frac{d\hat{\mathbf{x}}_\mathbf{t}}{dt} - \frac{1}{2}\mathbf{u}^T\frac{d\mathbf{P}_t}{dt}\mathbf{u}\right] \\ \frac{\partial\varphi}{\partial u} &= \varphi\left[i\hat{\mathbf{x}}_t - \mathbf{P}_t\mathbf{u}\right]\end{aligned} \tag{2.30}$$

Substituting these equations into Eq.(2.28) gives

$$\varphi i\mathbf{u}^T\frac{d\hat{\mathbf{x}}_\mathbf{t}}{dt} - \frac{1}{2}\varphi\mathbf{u}^T\frac{d\mathbf{P}_t}{dt}\mathbf{u} = \varphi i\mathbf{u}^T\mathbf{F}\hat{\mathbf{x}}_t - \frac{1}{2}\varphi\mathbf{u}^T\left(\mathbf{F}\mathbf{P}_t + \mathbf{P}_t\mathbf{F}^T\right)\mathbf{u} - \frac{1}{2}\varphi\mathbf{u}^T\mathbf{G}\mathbf{Q}\mathbf{G}^\mathbf{T}\mathbf{u} \tag{2.31}$$

By comparing the real and imaginary parts on both sides of the equation

$$\begin{aligned}\dot{\mathbf{x}}_\mathbf{t} &= \mathbf{F}(t)\mathbf{x}_t \\ \dot{\mathbf{P}}_t &= \left(\mathbf{F}(t)\mathbf{P}_t + \mathbf{P}_t\mathbf{F}^T(t)\right) - \mathbf{G}\mathbf{Q}\mathbf{G}^\mathbf{T}\end{aligned} \tag{2.32}$$

where $\dot{\mathbf{x}}_\mathbf{t} \triangleq \frac{d\hat{\mathbf{x}}_t}{dt}$ and $\dot{\mathbf{P}}_t \triangleq \frac{dP_t}{dt}$. These differential equations are the prediction equations for continuous-discrete Kalman filter, and the solutions of the equations are the prediction of the state $\hat{\mathbf{x}}_\mathbf{k}^-$ and the error covariance matrix $\mathbf{P}_k^-$. The measurement update equations can be found by identifying all three PDFs on the right side of Bayes' formula

$$p(\mathbf{x}, t_k|\mathbf{Y}_{t_k}) = \frac{p(\mathbf{y}_k|\mathbf{x})p(\mathbf{x}, t_k|\mathbf{Y}_{t_{k-1}})}{p(\mathbf{y}_k|\mathbf{Y}_{t_{k-1}})} \tag{2.33}$$

where $Y_{t_{k-1}} = [y_1, \ldots, y_{k-1}]$ is the history of the measurements. Since the measurement noise is assumed to be Gaussian, $p(\mathbf{y}_k|\mathbf{x}_k)$ can be expressed simply as

$$p(\mathbf{y}_k|\mathbf{x}_k) = \frac{1}{(2\pi)^{n/2}|\mathbf{R}|^{1/2}} e^{\{-\frac{1}{2}[\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k]^T\mathbf{R}^{-1}[\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k]\}} \tag{2.34}$$

From Eq. (2.34), the expectation of the measurement $\mathbf{y}_k$ is $E[\mathbf{y}_k|\mathbf{Y}_{t_{k-1}}] = \mathbf{H}_k\hat{\mathbf{x}}_k$, and the covariance can be calculated as

$$E\left[\left(\mathbf{y}_k - E[\mathbf{y}_k|\mathbf{Y}_{t_{k-1}}]\right)\left(\mathbf{y}_k - E[\mathbf{y}_k|\mathbf{Y}_{t_{k-1}}]\right)^T|\mathbf{Y}_{t_{k-1}}\right] = \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k \tag{2.35}$$

So the PDF of the measurement $\mathbf{y}_k$ given $\mathbf{Y}_{t_{k-1}}$ is simply

$$p(\mathbf{y}_k|\mathbf{Y}_{t_{k-1}}) \sim N(\mathbf{H}_k\hat{\mathbf{x}}_k,\ \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k) \tag{2.36}$$

Since all the PDFs in Bayes' formula are assumed to be Gaussian

$$p(\mathbf{x}, t_k|\mathbf{Y}_{t_{k-1}}) \sim N(\hat{\mathbf{x}}_k^-,\ \mathbf{P}_k^-) \tag{2.37}$$

By substituting equations (2.34), (2.36), and (2.37), Bayes' formula becomes

$$p(\mathbf{x}, t_k|\mathbf{Y}_{t_k}) = \frac{|\mathbf{H}\mathbf{P}^-\mathbf{H}^T|^{1/2}}{(2\pi)^{n/2}|\mathbf{R}|^{1/2}|\mathbf{P}^-|}\exp\left\{-\frac{1}{2}[*]\right\} \tag{2.38}$$

where

$$
\begin{aligned}
[*] = (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) + (\mathbf{x} - \hat{\mathbf{x}}^-)^T \left(\mathbf{P}^-\right)^{-1} (\mathbf{x} - \hat{\mathbf{x}}^-) \\
- (\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}^-)^T (\mathbf{H}\mathbf{P}^-\mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}^-)
\end{aligned}
\tag{2.39}
$$

All the subscripts and superscripts are omitted.

Since $p(\mathbf{x}, t_k|\mathbf{Y}_{t_k})$ is assumed to be Gaussian, i.e., $N(\hat{\mathbf{x}}_k^+,\ \mathbf{P}_k^+)$, the term $[*]$ must be equal to

$$
(\mathbf{x} - \hat{\mathbf{x}}^+)^T \left(\mathbf{P}^+\right)^{-1} (\mathbf{x} - \hat{\mathbf{x}}^+) \tag{2.40}
$$

Eq. (2.39) is rearranged to the form of the Eq. (2.40) as follows

$$
[*] = \mathit{\Gamma}^T \left[\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{P}^{-^{-1}}\right] \mathit{\Gamma} \tag{2.41}
$$

where $\mathit{\Gamma} = \left[\mathbf{x} - \left(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{P}^{-^{-1}}\right)^{-1} \left(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{y} + \mathbf{P}^{-^{-1}} \hat{\mathbf{x}}^{-1}\right)\right]$. By comparing Eq. (2.41) with Eq. (2.40) the update equations can be obtained as

$$
\hat{\mathbf{x}}^+ = \left(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{P}^{-^{-1}}\right)^{-1} \left(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{y} + \mathbf{P}^{-^{-1}} \hat{\mathbf{x}}^-\right) \tag{2.42}
$$

$$
\mathbf{P}^+ = \mathbf{P}^{-^{-1}} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \tag{2.43}
$$

They are not in the form in which the Kalman filter update equations are normally presented, and need to be reduced further. After matrix algebra Eq. (2.43) and Eq. (2.42) can be rewritten [38] as

$$
\hat{\mathbf{x}}^+ = \hat{\mathbf{x}}^- + \mathbf{P}^- \mathbf{H}^T \left(\mathbf{H}\mathbf{P}^-\mathbf{H}^T + \mathbf{R}\right)^{-1} \left(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}^-\right) \tag{2.44}
$$

$$
\mathbf{P}^+ = \mathbf{P}^- - \mathbf{P}^- \mathbf{H}^T \left[\mathbf{H}\mathbf{P}^-\mathbf{H}^T + \mathbf{R}\right]^{-1} \mathbf{H}\mathbf{P}^- \tag{2.45}
$$

The measurement update of the Kalman filter performs a correction using the measurement to estimate $\hat{\mathbf{x}}^+$ and the covariance $\mathbf{P}^+$ [53]. The equation (2.45) shows

that the update stage of the Kalman filter reduces the error covariance while it is increased in the prediction stage [69]. This means that more measurements make the estimation better. With the Kalman gain defined as $\mathbf{K} = \mathbf{P}^{-}\mathbf{H}^{T}\left(\mathbf{H}\mathbf{P}^{-}\mathbf{H}^{T} + \mathbf{R}\right)^{-1}$, the Kalman filter update equations can be presented in familiar form

$$\begin{aligned} \hat{\mathbf{x}}_k^{+} &= \hat{\mathbf{x}}_k^{-} + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_k^{-}\right) \\ \mathbf{P}_k^{+} &= \mathbf{P}_k^{-} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_k^{-} \\ \mathbf{K}_k &= \mathbf{P}_k^{-}\mathbf{H}_k^{T}\left(\mathbf{H}_k\mathbf{P}_k^{-}\mathbf{H}_k^{T} + \mathbf{R}_k\right)^{-1} \end{aligned} \tag{2.46}$$

The predictor-corrector structure of the Bayesian estimation in the Kalman filter is illustrated by Fig. 2.1.

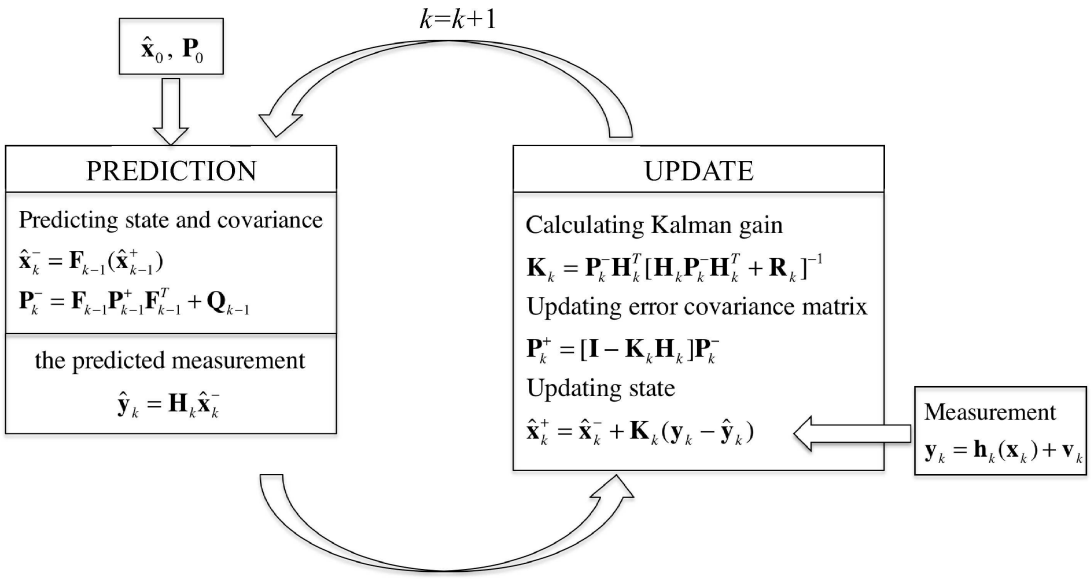


Figure 2.1: Flow Chart of the Kalman Filter

2.3 Nonlinear Filters

As shown in previous section, there exists an exact optimal solution for the linear filtering problem. However, most problems encountered in engineering are nonlinear (in system dynamics and/or measurement). So various techniques has been developed

for the nonlinear filtering problems, and some of them will be reviewed in this section.

2.3.1 Nonlinear Least Square Filter

The continuous form of the nonlinear stochastic equation for the system and the discrete form of the measurement equation are

$$\dot{\mathbf{x}}_t = \mathbf{f}(\mathbf{x}_t, t) + \mathbf{w}_t \tag{2.47}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \tag{2.48}$$

where the system noise $\mathbf{w_t}$ is a white Gaussian process $\mathbf{w_t} \sim N(0.Q_t)$, and measurement noise $\mathbf{v_k}$ are a white Gaussian sequence, $\mathbf{v_k} \sim N(0.R_k)$. and they are assumed to be independent each other. The description of the time function is simplified for convenience, for example, $\mathbf{x}_k \triangleq \mathbf{x}\,(t_k)$.

The nonlinear least square method is originally developed by Gaussian to determine planetary orbits [26]. The goal of this method is to find the estimation $\hat{\mathbf{x}}$ that minimizes the sum square of the residual errors.

$$J = \frac{1}{2}\mathbf{e}\mathbf{W}\mathbf{e}^T \tag{2.49}$$

where $\mathbf{e} = \mathbf{y} - \mathbf{h}(\mathbf{x})$. The measurement function $\mathbf{h}_k(\mathbf{x}_k)$ can be linearized about the estimation of the state $\hat{\mathbf{x}}$ by using the Taylor-series expansion as

$$\mathbf{h}(\mathbf{x}_k) = \mathbf{h}(\hat{\mathbf{x}}_k) + \left.\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}} (\mathbf{x}_k - \hat{\mathbf{x}}_k) + \text{H.O.T} \tag{2.50}$$

The current estimates of the state $\mathbf{x}$ are assumed to be

$$\mathbf{x}_c = [x_{1c}\ x_{2c} \cdots x_{nc}] \tag{2.51}$$

$\mathbf{x}_c$ and the estimates $\hat{\mathbf{x}}_k$ are assumed to be related by the unknown set of corrections $\Delta\mathbf{x}$ as

$$\hat{\mathbf{x}}_k = \mathbf{x}_c + \Delta\mathbf{x} \tag{2.52}$$

If the components of the corrections $\Delta\mathbf{x}$ are sufficiently small, it may be possible to solve for an approximation to the corrections and update $\mathbf{x}_c$ with an improved estimate of $\hat{\mathbf{x}}_k$ using the above equation. With this assumption, the function $\mathbf{h}(\hat{\mathbf{x}}_k, k)$ can be linearized about $\mathbf{x}_c$ using a first-order Taylor series expansion as

$$\mathbf{h}(\hat{\mathbf{x}}_k) \approx \mathbf{h}(\mathbf{x}_c) + \left.\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_c} \Delta\mathbf{x} \tag{2.53}$$

The measurement residual after the correction can be linearly approximated as

$$\Delta\mathbf{y} = \mathbf{y}_k - \hat{\mathbf{x}}_\mathbf{k} \approx \mathbf{y}_k - \mathbf{h}\mathbf{x}_c - \mathbf{H}\Delta\mathbf{x} = \Delta\mathbf{y}_c - \mathbf{H}\Delta\mathbf{x} \tag{2.54}$$

where $\mathbf{H} = \left.\frac{\partial \mathbf{h}}{\partial \mathbf{x}}\right|_{\mathbf{x}_c}$, and $\Delta\mathbf{x} = \hat{\mathbf{x}}_k - \mathbf{x}_c$. The approximation of the minimized weighted sum square J, given by Eq. (2.49) can be defined with $\Delta\mathbf{y}_c$ as

$$J_p = \frac{1}{2}\left(\Delta\mathbf{y}_c - \mathbf{H}\Delta\mathbf{x}\right)^T \mathbf{W}\left(\Delta\mathbf{y}_c - \mathbf{H}\Delta\mathbf{x}\right) \tag{2.55}$$

This is called the minimum sum of squares of the linearly predicted residuals, and minimizing J_p is equivalent to minimizing J, which is the original goal. If the process is convergent, $\Delta\mathbf{x}$ calculated by minimizing J_p should decrease until the linearization, given by Eq. (2.53) becomes a very good approximation of $\mathbf{h}(\hat{\mathbf{x}}_k)$ The minimization of J_p is basically same as solving the weighted least square problem. Therefore, it should satisfy the following conditions from the weighted least square problem

$$\bigtriangledown_{\Delta\mathbf{x}} J_p = \frac{\partial J_p}{\partial \Delta\mathbf{x}} = \mathbf{H}^T\mathbf{W}\mathbf{H}\Delta\mathbf{x} - \mathbf{H}^T\mathbf{W}\Delta\mathbf{y}_c = 0 \tag{2.56}$$

$$\nabla^2_{\Delta \mathbf{x}} J_p = \mathbf{H}^T \mathbf{W} \mathbf{H} > 0 \tag{2.57}$$

Eq. (2.56) and Eq. (2.57) are the necessary and the sufficient condition, respectively. From the necessary condition in Eq. (2.56), the correction $\Delta \mathbf{x}$ is obtained as

$$\Delta \mathbf{x} = \left(\mathbf{H}^T \mathbf{W} \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{W} \Delta \mathbf{y}_c \tag{2.58}$$

An initial guess of the current estimates $\mathbf{x}_c$ has to be made to begin the algorithm, and this process must be iterated until the very good approximation of $\mathbf{h}(\hat{\mathbf{x}}_k)$ is reached, i.e., $\Delta \mathbf{x}$ becomes near zero. A stopping condition can be made using the cost function J as

$$\delta J = \frac{|J_i - J_{i-1}|}{J_i} < \frac{\varepsilon}{\|\mathbf{W}\|} \tag{2.59}$$

where i is the iteration number and ε is a small number selected by user. The procedure will be iterated until the condition is met. However, it may have difficulty to converge when the $\mathbf{H}$ matrix has rank deficiency.

2.3.2 Extended Kalman Filter

The extended Kalman filter (EKF) delivers the minimum variance estimate of the state based on statistical information of the dynamic model and measurements. It employs the Kalman filter algorithm by linearizing nonlinear system dynamic models and/or nonlinear measurement models about the most recent estimate [27, 59]. The equations for the system (Eq. (2.47)) and the measurement (Eq. (2.48)) presented in previous section are also used here.

The EKF is based on the Taylor series expansion of the nonlinear systems and/or measurement equations about the current estimated value $\hat{\mathbf{x}}_k$. Thus, the prediction of the state estimation and covariance are done by [27]

$$\hat{\mathbf{x}}_k^- = \mathbf{f}(\hat{\mathbf{x}}_{k-1}^+) \tag{2.60}$$

$$\dot{\mathbf{P}}_t = \mathbf{\Phi}_{k-1}\mathbf{P}_t\mathbf{\Phi}_{k-1}^T + \mathbf{Q}_t \tag{2.61}$$

where $\mathbf{\Phi}_k \approx \partial\mathbf{f}(\mathbf{x}_t)/\partial\mathbf{x}|_{\mathbf{x}_t=\hat{\mathbf{x}}_{k-1}^+}$ is the Jacobian matrix of the nonlinear system model $\mathbf{f}(\cdot)$ evaluated around the current state. The update equations for the state and the covariance are

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k\left[\mathbf{y}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-)\right] \tag{2.62}$$

$$\mathbf{P}_k^+ = [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\mathbf{P}_k^- \tag{2.63}$$

where $\mathbf{H}_k \approx \partial\mathbf{h}_k/\partial\mathbf{x}|_{\mathbf{x}=\hat{\mathbf{x}}_k^-}$ is the Jacobian matrix of the nonlinear measurement model $\mathbf{h}_k(\cdot)$ evaluated about the predicted state $\mathbf{x} = \hat{\mathbf{x}}_k^-$, and the Kalman gain is found by

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^T[\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^T + \mathbf{R}_k]^{-1} \tag{2.64}$$

If a large uncertainty is involved in the estimation the above error covariance update equation can lead to filter instability due to the inversion of the error covariance on the right side of Eq. (2.63). This can be avoided using the *Joseph form* of the error covariance update equation [13]

$$\mathbf{P}_k^+ = [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\mathbf{P}_k^-[\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \tag{2.65}$$

The EKF assumes the state distribution as Gaussian, and the state has gone through the first-order linearization of the nonlinear functions. These have been pointed out as the sources of the large errors in the true mean and covariance [5, 32, 35, 40]. In the next section, a new approach for approximating the mean and covariance without the first-order linearization of nonlinear functions will be discussed.

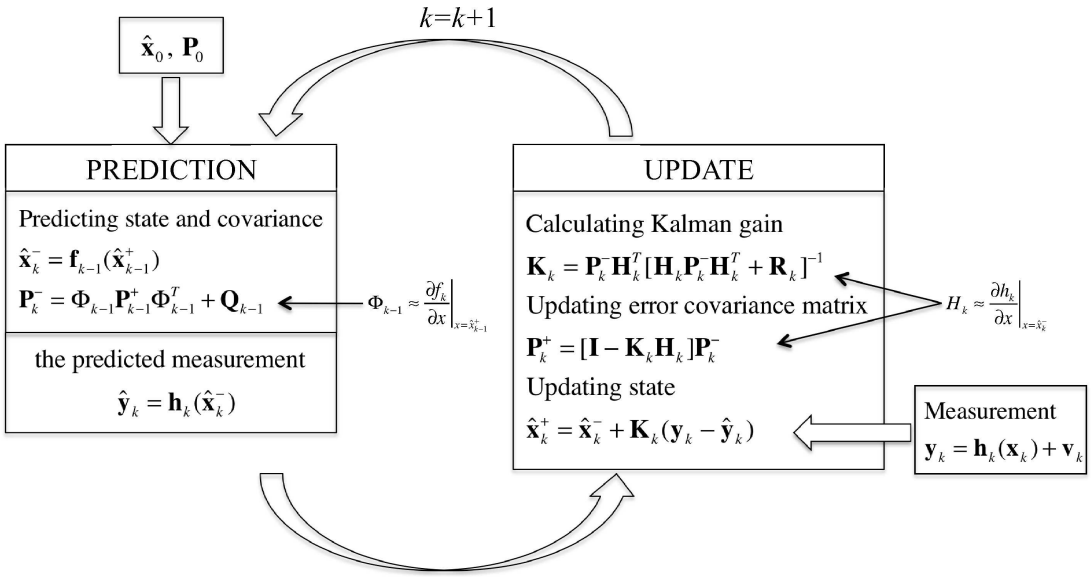


Figure 2.2: Flow Chart of the Extended Kalman Filter

2.3.3 Unscented Kalman Filter

The most popular estimator for nonlinear system is probably the extended Kalman filter [33]. The Extended Kalman filter employs the Kalman filter algorithm by linearizing the nonlinear system dynamic and/or measurement models, which makes the concept of the EKF straightforward but causes the following drawbacks [40, 43].

- When the system and/or observation models are highly non-linear, the extended Kalman filter can give a poor performance as a consequence of the linearization.
- The derivation of the Jacobian is not a trivial task in many applications.
- Numerical evaluation of the Jacobian can be computationally intensive.

In the UKF, the nonlinear transformation of a random variable is done through the unscented transformation (UT). The result is the filter that can capture the true mean and covariance more accurately. In addition, this technique removes the need to calculate Jacobian, which can be a difficult task for complex functions.

The unscented transformation is a method for calculating the statistics of a random variable, which undergoes a nonlinear transformation. It is built on the idea

that it is easier to approximate a probability distribution than an arbitrary nonlinear function [42, 43]. The UT utilizes a set of points called sigma points to represent the mean and the covariance accurately up to the second order of the Taylor series expansion for an arbitrary nonlinear function [41]. Although this method seems similar to Monte Carlo-type methods, there is a big difference in the basic principle. Unlike Monte Carlo-type methods the UT selects the samples, i.e., sigma points, based on mean and covariance by a specific deterministic algorithm.[40]. There are several different kinds of UT techniques available, depending on how the sigma points are chosen and how many of them are used [32, 39]. In this section the scaled unscented transformation (SUT) [42] and the UKF based on the SUT will be presented [35, 42]. A comparison between the UT and the linearization approach is shown in Fig. 2.3. The upper figure shows the mean and covariance obtained from the UT, whereas the lower figure shows the results from a linearization approach.

Scaled Unscented Transformation

Considering the propagation of a random variable $\mathbf{x}$ with a mean value of $\bar{x}$ and a covariance matrix of $\mathbf{P}_{xx}$ through a nonlinear function $\mathbf{y} = \mathbf{f}(\mathbf{x})$. To calculate the statistics of $\mathbf{y}$, a new matrix $\boldsymbol{\mathcal{X}}$ that consists of $(2n+1)$ weighted sigma vectors $\mathcal{X}_i$ is formed according to the following

$$\begin{aligned} \mathcal{X}_0 &= \bar{\mathbf{x}} \\ \mathcal{X}_i &= \bar{\mathbf{x}} + \left(\sqrt{(n+\lambda)\,\mathbf{P}_{xx}}\right)_i \qquad i = 1, \ldots, n \\ \mathcal{X}_i &= \bar{\mathbf{x}} - \left(\sqrt{(n+\lambda)\,\mathbf{P}_{xx}}\right)_{i-n} \qquad i = n+1, \ldots, 2n \end{aligned} \tag{2.66}$$

These sigma points are propagated through the nonlinear function

$$\mathcal{Y}_i = \mathbf{f}(\mathcal{X}_i) \qquad i = 0, \ldots, 2n \tag{2.67}$$

The estimated mean and covariance of $\mathbf{y}$ are approximated using a weighted sample mean and covariance of the posterior sigma points, respectively as

$$\bar{\mathbf{y}} \approx \sum_{i=0}^{2n} W_i^m \mathcal{Y}_i \tag{2.68}$$

$$\mathbf{P}_{yy} \approx \sum_{i=0}^{2n} W_i^c (\mathcal{Y}_i - \bar{\mathbf{y}})(\mathcal{Y}_i - \bar{\mathbf{y}})^T \tag{2.69}$$

The associated scalar weights W_i^m and W_i^c are calculated as

$$\begin{aligned} W_0^m &= \frac{\lambda}{n+\lambda} \\ W_0^c &= \frac{\lambda}{n+\lambda} + \left(1 - \alpha^2 + \beta\right) \\ W_i^m &= W_i^c = \frac{1}{2\left(n+\lambda\right)} \qquad i = 1, 2 \cdots 2n \end{aligned} \tag{2.70}$$

where $\lambda = \alpha^2(n+\kappa) - n$ is a scaling parameter. The constant α control the distance between the sigma points and the mean $\bar{\mathbf{x}}$, which is well illustrated in Refs. [32, 42], and is usually chosen to be a small positive number (e.g. $1 \leq \alpha \leq 10^{-4}$). The constant κ is another scaling parameter that provides an extra degree of freedom to fine tune the higher order moments of the approximation, and $\kappa = 3 - n$ is usually used [35]. When $\kappa = 0$, the first sigma point $\mathcal{X}_0$ is effectively excluded [32]. β is the third parameter that incorporates further higher order effects by adding the weight on the zeroth sigma point of the calculation of the covariance, and $\beta = 2$ is optimal for Gaussian distributions [35, 42].

Unscented Kalman Filter

The UKF is a straightforward extension of the UT to the recursive estimation. The UKF is derived for discrete-time nonlinear equations presented as

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{w}_k, k) \tag{2.71}$$

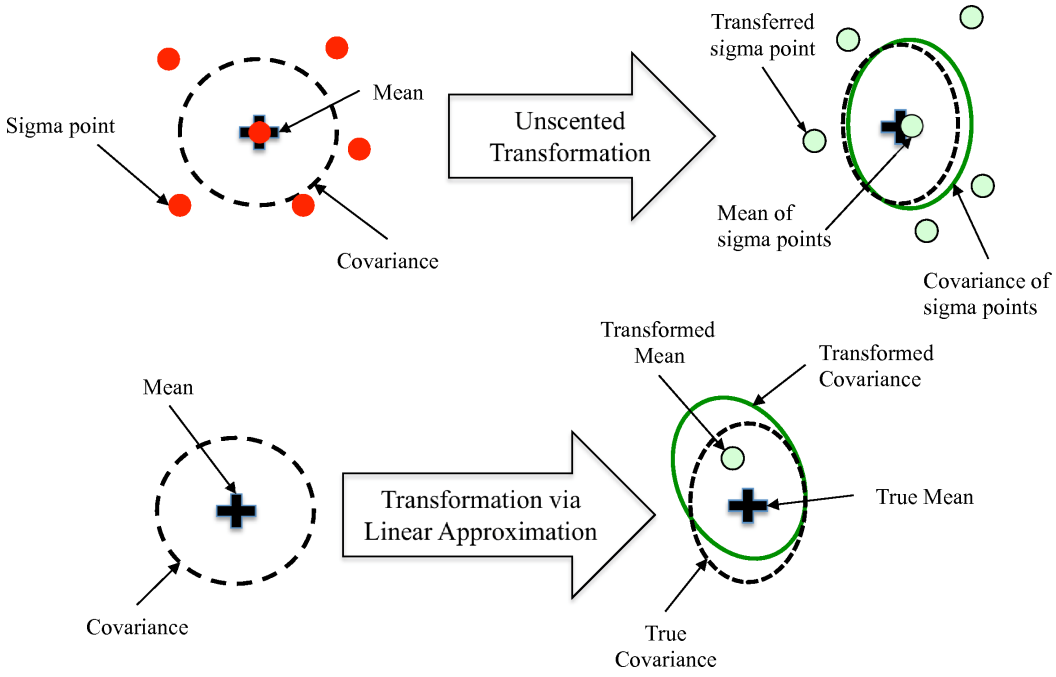


Figure 2.3: Illustration of the Unscented Transformation

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{v}_k, k) \tag{2.72}$$

where $\mathbf{x}_k \in \mathbb{R}^{n\times 1}$ is the state vector and $y_k \in \mathbb{R}^{m\times 1}$ is the measurement vector at time k. $\mathbf{w}_k \in \mathbb{R}^{q\times 1}$ is a process noise vector and $\mathbf{v}_k \in \mathbb{R}^{r\times 1}$ is the additive measurement noise vector, and they are assumed to be $N(0, \mathbf{Q}_k)$ and $N(0, \mathbf{R}_k)$, respectively. The original state vector is redefined as an augmented state vector along with process noise variables as $\mathbf{x}_k^a = \begin{bmatrix}\mathbf{x}_k^T & \mathbf{v}_k^T & \mathbf{w}_k^T\end{bmatrix}^T$. An augmented covariance matrix is constructed as

$$\mathbf{P}_k^a = \begin{bmatrix} \mathbf{P}_k & \mathbf{P}_k^{xw} & \mathbf{P}_k^{xv} \\ (\mathbf{P}_k^{xw})^T & \mathbf{Q}_k & \mathbf{P}_k^{wv} \\ (\mathbf{P}_k^{xv})^T & (\mathbf{P}_k^{wv})^T & \mathbf{R}_k \end{bmatrix} \tag{2.73}$$

where $\mathbf{P}_k^{xw}$ and $\mathbf{P}_k^{xv}$ are the correlation between the error in the state estimation and the process noise and the error between the state estimation and the measurement noise, and $\mathbf{P}_k^{wv}$ is the correlation between the process noise and the measurement noise, which are all zero for most systems. The sigma points of the augmented state

vector can be found by Eq. (2.66), and the number of sigma points for the augmented state vector is $2(n+m+q)+1 \triangleq L+1$. The state propagation, the predicted state vector and its predicted covariance can be obtained by applying the scaled unscented transformations and following the EKF algorithm as

$$\mathcal{X}^a_{i,k+1} = \mathbf{f}(\mathcal{X}^a_{i,k}, k) \tag{2.74}$$

$$\hat{\mathbf{x}}^-_{k+1} = \sum_{i=0}^{2L} W_i^m \mathcal{X}^a_{i,k+1} \tag{2.75}$$

$$\mathbf{P}^-_{k+1} = \sum_{i=0}^{2L} W_i^c \left(\mathcal{X}^a_{i,k+1} - \hat{\mathbf{x}}^-_{k+1}\right)\left(\mathcal{X}^a_{i,k+1} - \hat{\mathbf{x}}^-_{k+1}\right)^T \tag{2.76}$$

The predicted measurement and its predicted covariance are also calculated as

$$\mathcal{Y}_{i,k+1} = h(\mathcal{X}^a_{i,k+1}, k+1) \tag{2.77}$$

$$\hat{\mathbf{y}}^-_{k+1} = \sum_{i=0}^{2L} W_i^m \mathcal{Y}_{i,k+1} \tag{2.78}$$

$$\mathbf{P}^{yy}_{k+1} = \sum_{i=0}^{2L} W_i^c \left(\mathcal{Y}_{i,k+1} - \hat{\mathbf{y}}^-_{k+1}\right)\left(\mathcal{Y}_{i,k+1} - \hat{\mathbf{y}}^-_{k+1}\right)^T \tag{2.79}$$

The Kalman gain is $\boldsymbol{\mathcal{K}}_{k+1} = \mathbf{P}^{xy}_{k+1}\mathbf{P}^{yy}_{k+1}$ where the cross covariance can be determined by

$$\mathbf{P}^{xy}_{k+1} = \sum_{i=0}^{2L} W_i^c \left(\mathcal{X}^a_{i,k+1} - \hat{\mathbf{x}}^-_{k+1}\right)\left(\mathcal{Y}_{i,k+1} - \hat{\mathbf{y}}^-_{k+1}\right)^T \tag{2.80}$$

Finally, the estimation of the state and the error covariance are given as

$$\hat{\mathbf{x}}^+_{k+1} = \hat{\mathbf{x}}^-_{k+1} + \boldsymbol{\mathcal{K}}_{k+1}(\mathbf{y}_{k+1} - \hat{\mathbf{y}}_{k+1}) \tag{2.81}$$

$$\mathbf{P}^+_k = \mathbf{P}^-_k - \boldsymbol{\mathcal{K}}_{k+1}\mathbf{P}^{yy}_{k+1}\boldsymbol{\mathcal{K}}^T_{k+1} \tag{2.82}$$

No explicit calculation of the Jacobian and/or Hessian matrix is needed to implement this algorithm, and the formulation is ideally suited for parallel computation since

the propagations can be performed in parallel. Note that the process model and the measurement model is in a more general form than those in previous section. In many cases, the process and the measurement noise are purely additive. With purely additive noise the computational complexity of the UKF can be reduced by not augmenting the state vector and the error covariance with the noise and a set of equations without augmentation is presented in Ref. [35], which is claimed to produce the same result as the augmented equations presented in this section. The comparison study of the augmented and the non-augmented approach has been done and found that it is necessary to meet certain condition, i.e. $\kappa = 3 - n$, for these two approaches to become equivalent [74].

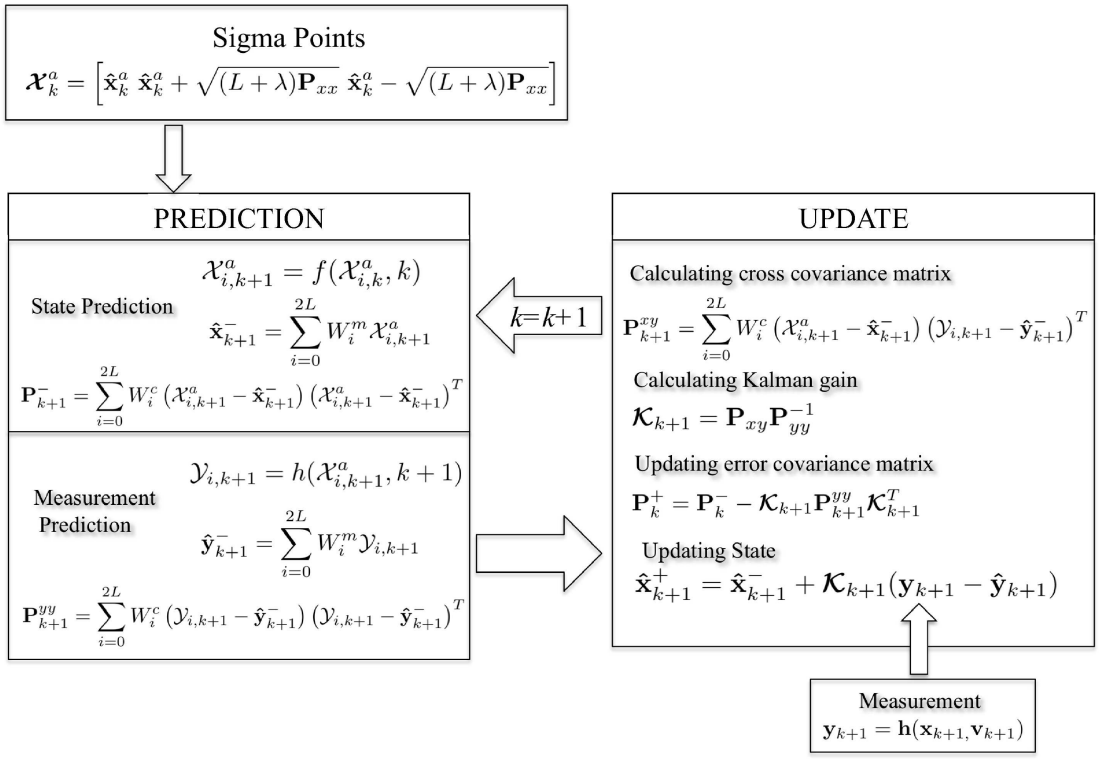


Figure 2.4: Flow Chart of the Unscented Kalman Filter

2.3.4 Particle Filter

The basic idea of Sequential Monte Carlo (SMC) in the form of important sampling was introduced in statistics in the 1950s [34], but was mostly overlooked and ignored due to the lack of computer power and the degeneration of the particle over time. The major breakthrough of the SMC method was the development of the resampling technique that solved the degeneration problem [31]. Since then, the interest in this method has dramatically increased [22] generating many improvements of particle filters. SMC techniques achieve the filtering based on a set of weighted point mass (or "particle") of the state variables. The particles and their corresponding weights are combined to represent the posterior PDF. After a certain number of recursive steps most of the weights of particles will become negligible, i.e., close to zero. This is the degeneracy phenomenon. To prevent this phenomenon, the particle with smaller weights will be removed and those with larger weights multiplied based on some kind of evolution process. Therefore, the implementation of the particle filter consists of three important steps; 1) generation of particles, 2) computation of the particle weights, and 3) resampling. The theory of particle filtering is presented in a brief manner without proofs. More details and an investigation of particle filters are available in Ref. [65].

Sequential Important Sampling

Consider the nonlinear state space model

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{w}_k) \tag{2.83}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \tag{2.84}$$

where $\mathbf{x}_k \in \mathbb{R}^{n\times 1}$ is the state vector and $y_k \in \mathbb{R}^{m\times 1}$ is the measurement vector at time k. $\mathbf{w}_k \in \mathbb{R}^{q\times 1}$ is a process noise vector and $\mathbf{v}_k \in \mathbb{R}^{r\times 1}$ is the additive measurement noise vector, and they are assumed to be $N(0, \mathbf{Q}_k)$ and $N(0, \mathbf{R}_k)$, respectively. If the system is Markovian, which means $p(\mathbf{x}_i|\mathbf{x}_{i-1}) = p(\mathbf{x}_i|\mathbf{x}_{i-1}, \cdots, \mathbf{x}_0)$, then the state PDF and the measurement PDF can be expressed as

$$p(\mathbf{X}_k) = p(\mathbf{x}_0)\prod_{i=1}^{k} p(\mathbf{x}_i|\mathbf{x}_{i-1}) \tag{2.85}$$

and

$$p(\mathbf{Y}_k|\mathbf{X}_k) = \prod_{i=0}^{k} p(\mathbf{y}_i|\mathbf{x}_i) \tag{2.86}$$

where $\mathbf{X}_k = [\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k]$ and $\mathbf{Y}_k = [\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_k]$. The posterior PDF in terms of $p(\mathbf{X}_k|\mathbf{X}_k, \mathbf{Y}_{k-1})$, $p(\mathbf{y}_k|\mathbf{x}_k)$, and $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ can be expressed in form of a recursive equation.

$$\begin{aligned} p(\mathbf{X}_k|\mathbf{Y}_k) &= \frac{p(\mathbf{y}_k|\mathbf{X}_k, \mathbf{Y}_{k-1})p(\mathbf{X}_k|\mathbf{Y}_{k-1})}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})} \\ &= \frac{p(\mathbf{y}_k|\mathbf{X}_k, \mathbf{Y}_{k-1})p(\mathbf{x}_k|\mathbf{X}_{k-1}, \mathbf{Y}_{k-1})p(\mathbf{X}_k|\mathbf{Y}_{k-1})}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})} \\ &= \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})}{p(\mathbf{y}_k|\mathbf{Y}_{k-1})}p(\mathbf{X}_{k-1}|\mathbf{Y}_{k-1}) \\ &\propto p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{x}_{k-1})p(\mathbf{X}_{k-1}|\mathbf{Y}_{k-1}) \end{aligned} \tag{2.87}$$

The posterior PDF $p(\mathbf{x}_k|\mathbf{Y}_k)$ from above equation can be approximated using a large number of samples, i.e., particles.

$$\hat{p}(\mathbf{X}_k|\mathbf{Y}_k) = \frac{1}{N}\sum_{i=1}^{N} \delta\left(\mathbf{X}_k - \mathbf{X}_k^i\right) \tag{2.88}$$

where $\{\mathbf{X}_k^i\}_{i=1}^N$ is the samples selected form the posterior, and $\delta(\mathbf{X}_k)$ is the Dirac delta function. Since the samples are collected from the posterior itself, the weights are equal, and their sum is one. However, the samples cannot be drawn from the

posterior since it is unknown for most cases. Instead, the samples are drawn from a known PDF $q(\mathbf{X}_k|\mathbf{Y}_k)$ called importance density, and the weights can be presented in the following relation

$$w_k^i \propto \frac{p(\mathbf{X}_k^i|\mathbf{Y}_k)}{q(\mathbf{X}_k^i|\mathbf{Y}_k)} = \frac{p(\mathbf{y}_k|\mathbf{x}_k^1)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)p(\mathbf{X}_{k-1}^i|\mathbf{Y}_{k-1})}{q(\mathbf{x}_k^i|\mathbf{X}_{k-1}^i,\mathbf{Y}_k)q(\mathbf{X}_{k-1}^i|\mathbf{Y}_{k-1})} \tag{2.89}$$

and

$$w_k^i = w_{k-1}^i \frac{p(\mathbf{y}_k|\mathbf{x}_k^1)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)}{q(\mathbf{x}_k^i|\mathbf{X}_{k-1}^i,\mathbf{Y}_k)} \tag{2.90}$$

where $q(\mathbf{x}_k^i|\mathbf{X}_{k-1}^i,\mathbf{Y}_k)q(\mathbf{X}_{k-1}^i|\mathbf{Y}_{k-1})$ is from the factorization of $q(\mathbf{X}_k^i|\mathbf{Y}_k)$ If the importance density is chosen such that $q(\mathbf{x}_k^i|\mathbf{X}_{k-1}^i,\mathbf{Y}_k) = q(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i,\mathbf{Y}_k)$, the weight becomes

$$w_k^i = w_{k-1}^i \frac{p(\mathbf{y}_k|\mathbf{x}_k^1)p(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i)}{q(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i,\mathbf{Y}_k)} \tag{2.91}$$

and

$$\hat{p}(\mathbf{x}_k|\mathbf{Y}_k) = \sum_{i=1}^{N} w_k^i \delta\left(\mathbf{x}_k - \mathbf{x}_k^i\right) \tag{2.92}$$

The estimation of $\mathbf{x}$ and the covariance can be obtained by the following integral

$$\begin{aligned} \hat{\mathbf{x}}_k^+ &= \int \mathbf{x}_k \hat{p}(\mathbf{x}_k|\mathbf{Y}_k) d\mathbf{x}_k = \int \mathbf{x}_k \sum_{i=1}^{N} w_k^i \delta d\mathbf{x}_k \\ &= \sum_{i=1}^{N} w_k^i \mathbf{x}_k^i \end{aligned} \tag{2.93}$$

and

$$\begin{aligned} \mathbf{P}_k^+ &= \int \left(\mathbf{x}_k - \hat{\mathbf{x}}_k^+\right)\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^+\right)^T \hat{p}(\mathbf{x}_k|\mathbf{Y}_k) d\mathbf{x}_k \\ &= \sum_{i=1}^{N} w_k^i \left(\mathbf{x}_k^i - \hat{\mathbf{x}}_k^+\right)\left(\mathbf{x}_k^i - \hat{\mathbf{x}}_k^+\right)^T \end{aligned} \tag{2.94}$$

2.3.5 Resampling

The sample are drawn from the importance density function though it is ideal that the samples are collected from the posterior PDF it self. It has been shown [21] that the variance of the importance weights can only increase over time with the importance density function. The variance increase has a damaging effect on the accuracy of the filter and leads to the problem known as the degeneracy phenomena. This means that after a certain number of recursive processes, most of the particles will have very small weights, and it is not avoidable. This was the major hurdle in the development of the particle filter. The effective sample size N_{eff} as a measure of degeneracy of an algorithm was introduced [47] as

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N} (w_k^i)^2} \tag{2.95}$$

If all weights are uniform, i.e., $w_k^i = \frac{1}{N}$, the effective sample size will be N while $\hat{N}_{eff} = 1$ if all but one particle have zero weights. This is an indicator of how well the particles are concentrated in the area that contributes more to the approximation of $\hat{p}(\mathbf{x}_k|\mathbf{Y}_k)$. A lower threshold can be set, so when the $\hat{N}_{eff}$ falls below the threshold, N new samples can be selected, i.e. resampling. Resampling removes the samples with low weights and increases the number of samples with high weights by drawing the new samples form the estimated posterior and the weight are set to be $1/N$. After resampling, the samples are not independent any more since part of samples are the copies of the same sample. Therefore, the number of resamplings should be kept at aminimum.

2.3.6 Resampling Methods

The basic idea of resampling is to eliminate particles with small weights and to concentrate on particle with large weights. The simplest resampling method would be a direct approach implementing the resampling that generates N independent and identically distributed (i.i.d.) random variables from the uniform distribution, sorts them in ascending order and compares them with the cumulative sum of normalized weights [65]. First,calculate the thresholds using a cumulative sum of the normalized weights in any order. Then, for each index i

1. Draw a uniform random number u_i based on the uniform distribution.

2. Use a search algorithm (binary search) to locate the position of u_i within the thresholds

3. Set the resampled index according to the index of the location of u_i

The idea of the random resampling algorithm is simple, but it is computationally inefficient due to the fact that the best sorting algorithm has a complexity of $O(NlogN)$. An algorithm based on order statistics [14, 64] can be implemented with a complexity of $O(N)$. There other efficient (in terms of reducing variation) resampling methods such as stratified sampling, residual sampling [54], systematic resampling [46], etc. The detail description of the each method is omitted since the theory described in this chapter is only to be considered as a brief description of how particle filters work

CHAPTER 3

Nonlinear Filtering via Numerical Solution of Fokker-Planck Equation

In this chapter, a direct numerical approximation to the optimal nonlinear filter is investigated. The exact nonlinear filter for systems with continuous nonlinear dynamics and discrete nonlinear measurements consists of two equations (Fig. 3.1) [38]. A partial differential equation called the Fokker-Planck equation (FPE) [24, 62] describes how the conditional density evolves between measurements, and Bayes' formula describes how the conditional density is modified using measurements. In Daum's paper [19] where the characteristics of difference types of nonlinear filtering were described, he points out that the state estimation of this type of algorithm can be optimal if designed carefully.

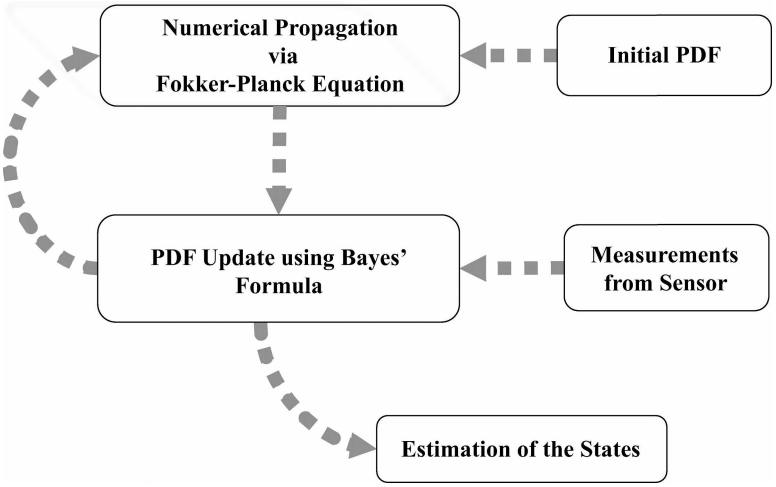


Figure 3.1: Nonlinear Estimation via FPE

3.1 Fokker-Planck Equation and Nonlinear Filtering

The Itô equation describes how the states and their probability densities of a dynamic system evolve in time. For the state $\mathbf{x}_t \in \mathbb{R}^{n\times 1}$, the Itô stochastic differential equation is

$$d\mathbf{x}_t = \mathbf{f}(x_t, t)dt + \mathbf{G}(x_t, t)d\boldsymbol{\beta}_t, \quad t \geq t_0 \tag{3.1}$$

where $\mathbf{f}(t) \in \mathbb{R}^{n\times 1}$ is a dynamic model, $\mathbf{x}_t \in \mathbb{R}^{n\times 1}$ is the state vector, $\boldsymbol{\beta}_t$ is an m-vector Brownian motion process with $E[\boldsymbol{\beta}_t\boldsymbol{\beta}_t^T] = \mathbf{Q}(t)dt$, and $\mathbf{G}(t) \in \mathbb{R}^{n\times m}$ is a matrix function. In the Itô equation, $\mathbf{f}$ models the deterministic part of the dynamics while G represents the random part of the dynamics. The discrete nonlinear measurement y_k taken at discrete time instants t_k is

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_{t_k}, t_k) + \mathbf{v}_k, \quad k = 1, 2, ... \text{ and } t_{k+1} > t_k \geq t_0 \tag{3.2}$$

where $\mathbf{h}_k \in \mathbb{R}^{m\times 1}$ is a measurement model, $\mathbf{y}_k \in \mathbb{R}^{m\times 1}$ is the measurement vector, and $\mathbf{v}_k$ is a white Gaussian noise with $N(0, R_k)$.

Assuming *apriori* information $p(\mathbf{x}, t_0|\mathbf{Y}_{t_0}) = p(\mathbf{x}_{t_0})$ is given and $\mathbf{Y}_t = [\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_k]$ is the set of measurements taken up to time t_k, the solution of the filtering problem $p(\mathbf{x}, t|\mathbf{Y}_t)$ can be determined by the following theorem by Jazwinski [38].

Theorem. *Suppose the prior density $p(\mathbf{x}, t)$ for Eq. (3.1) exists and is once continuously differentiable with respect to t and twice with respect to x. Let h be continuous in both arguments and bounded for each t_k wp 1. Then between measurements, i.e., $t_k \leq t < t_{k+1}$, the conditional density $p(\mathbf{x}, t|\mathbf{Y}_t)$ satisfies Fokker-Planck/Kolmogorov's forward equation.*

When a measurement is made at t_k, the conditional density satisfies the difference equation

$$p(\mathbf{x}, t_k|\mathbf{Y}_{t_k}) = \frac{p(\mathbf{x}, t_k|\mathbf{Y}_{t_{k-1}})p(\mathbf{y}_k|\mathbf{x})}{\int p(\xi|\mathbf{Y}_{t_{k-1}})p(\mathbf{y}_t|\xi)d\xi} \tag{3.3}$$

where $p(\mathbf{y}_k|\mathbf{x})$ *is given by*

$$p(\mathbf{y}_k|\mathbf{x}) = \frac{1}{(2\pi)^{m/2}|R_k|^{1/2}} e^{\left\{-\frac{1}{2}\left[\mathbf{y}_k - h(\mathbf{x},t_k)\right]^T R_k^{-1}\left[\mathbf{y}_k - h(\mathbf{x},t_k)\right]\right\}} \tag{3.4}$$

Prediction $p(\mathbf{x}, t_k|\mathbf{Y}_{t_{k-1}})$ *is accomplished via Fokker-Planck/Kolmogorov's forward equation.*

Once $p(\mathbf{x}, t_k|\mathbf{Y}_{t_k})$ is obtained, the optimal estimate in the sense of minimum mean square error (MMSE) can be obtained by the conditional expectation of $\mathbf{x}_k$ given $\mathbf{Y}_k$

$$\hat{\mathbf{x}}_k = E[\mathbf{x}|\mathbf{Y}_k] = \int \mathbf{x} p(\mathbf{x}, t_k|\mathbf{Y}_{t_k}) d\mathbf{x} \tag{3.5}$$

The Fokker-Planck equation (FPE), which is also known as the Kolmogorov forward equation, is first used by Fokker [24] and Planck [24] to explain the Brownian motion of particles. The equation can explain the behavior of a dynamic system that depicts the characteristic of the Brownian motion. Stochastic systems can be modeled as an n-dimensional Itô stochastic differential equation (SDE), given by Eq. (3.1). If the process described by the SDE is a Markovian process, the probability density function characterizing this process is governed by the Fokker-Planck equation.

$$\frac{\partial p}{\partial t} = -\sum_{i=1}^{n} \frac{\partial \left[p\mathbf{f}_i\right]}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 \left[p\left(\mathbf{G}\mathbf{Q}\mathbf{G}^T\right)_{ij}\right]}{\partial x_i \partial x_j} \tag{3.6}$$

where p is the conditional probability density function $p\left(\mathbf{x}, t|\, \mathbf{Y}_k\right)$. The FPE is a partial differential equation describing the time evolution of the probability density function of the Markov process governed by the Itô SDE. The equation is often called the equation of the motion of the probability density. The first term on the right side of the FPE is called the drift term, responsible for the evolution of the state of the system. The second term is the diffusion, responsible for the statistical evolution of the system. The equation becomes the equation of deterministic process if the

diffusion term is neglected.

For the filter based on the solution of the FPE to work properly, solving the FPE quickly and accurately is very crucial. The analytical solution of FPE is difficult to be obtained with a few exceptions. Therefore the evaluation of FPE between measurements usually has to be done numerically. This has prevented the utilization of the nonlinear filtering based on the solution of FPE until recent years. Due to the complexity in finding analytical solutions and the usefulness of FPE in modern physics, quite a few numerical schemes have been developed [10, 25] such as the finite difference method [23], path-integral method[1], cell-mapping method [63], the method based on distributed approximating functional [81], and the finite element method [50, 51]

3.2 Filter Based on the Finite Difference Methods

3.2.1 Explicit Forward Method

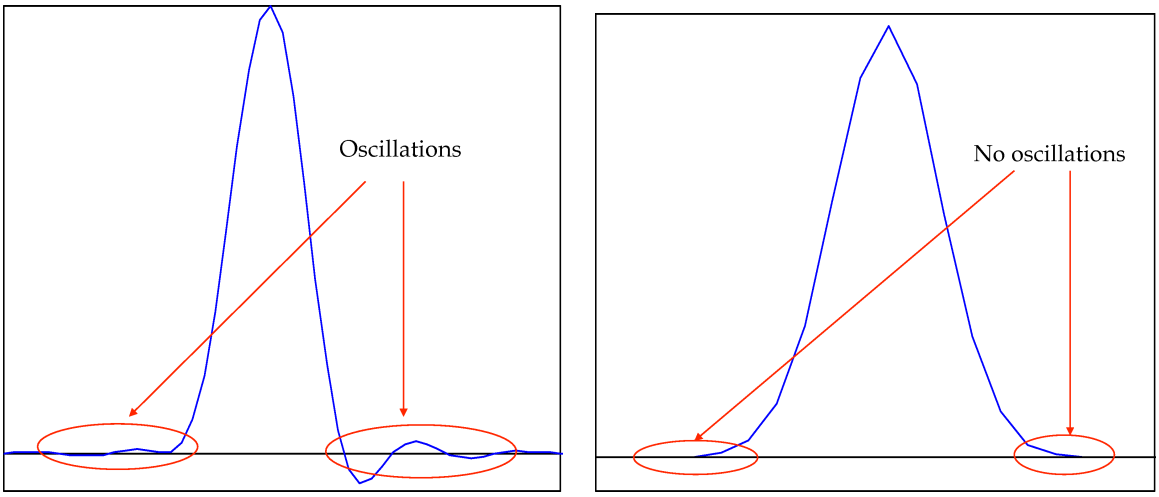
Since it is difficult to solve Eq. (3.6) analytically, use of a numerical method is mostly unavoidable as mentioned earlier. First, both the MacCormack method [4] and a simple explicit finite difference method with central differencing in space were tried. Then, it was determined to use the explicit finite difference method for its speed and simplicity.

$$\begin{aligned}
\left.\frac{\partial p\left(x,t\right)}{\partial t}\right|_{t=t_i} &= \frac{p\left(x,t_i\right)-p\left(x,t_{i-1}\right)}{\Delta t} \\
\left.\frac{\partial p\left(x,t\right)}{\partial X}\right|_{t=t_{i-1}} &= \frac{p\left(x_{j+1},t_{i-1}\right)-p\left(x_{j-1},t_{i-1}\right)}{2\Delta x} \\
\left.\frac{\partial^2 p\left(x,t\right)}{\partial X^2}\right|_{t=t_{i-1}} &= \frac{p\left(x_{j+1},t_{i-1}\right)-2p\left(x_j,t_{i-1}\right)+p\left(X_{j-1},t_{i-1}\right)}{\Delta x^2}
\end{aligned} \tag{3.7}$$

where $\Delta t = t_i - t_{i-1},\ i = 1, 2 \ldots$ and $\Delta x = x_j - x_{j-1},\ j = 1, 2 \ldots$ is grid spacing When the central differencing is applied to Eq. (3.6), it can be seen that spacial oscillations are generated at both the leading and trailing edges of the PDF. It turns out that all central differencing schemes for solving the advective equation suffer from a similar problem as shown in Fig. 3.2(a) To suppress the oscillations as shown in Fig. 3.2(b), upwind differencing scheme [4, 70] was employed for all of the first order spatial derivative terms as

$$\left.\frac{\partial p\left(X,t\right)}{\partial X}\right|_{t=t_{i-1}} = a \begin{cases} \dfrac{p\left(x_{j+1},t_{i-1}\right) - p\left(x_j,t_{i-1}\right)}{\Delta x}, & a \geq 0 \\ \dfrac{p\left(x_j,t_{i-1}\right) - p\left(x_{j-1},t_{i-1}\right)}{\Delta x}, & a < 0 \end{cases} \tag{3.8}$$

The accuracies of the central differencing method are $O\left(\Delta x^2, \Delta y^2, \Delta z^2\right)$, and the



(a) Central differencing scheme (b) Upwind differencing schems

Figure 3.2: Central differencing vs. Upwind differencing

accuracies of the upwind differencing and the explicit forward differencing in time are $O\left(\Delta x, \Delta y, \Delta z\right)$ and $O\left(\Delta t\right)$ respectively. So the overall accuracy of the numerical approximation is $O\left(\Delta x, \Delta y, \Delta z, \Delta t\right)$. For this explicit finite difference scheme to be

stable, it was found that the following inequality must be satisfied [4].

$$0 < \Delta t \left(\frac{1}{\Delta x_1^2} + \frac{1}{\Delta x_2^2} \cdots \right) < \frac{1}{2} \tag{3.9}$$

where $\Delta x_1, \Delta x_2, \ldots$ are the grid size in different directions. The boundary conditions are assumed to satisfy the homogeneous Dirichlet conditions [15] that the boundary values are kept at zero since a PDF decays exponentially and the domain is large enough that the probability density at the boundary is close to zero.

Because of the stability condition presented above, the time step has to be very small for the problem that requires very small grid spacing. This can cause a significant rise in the computational cost for even low dimensional problems, and calls for the implicit method that is unconditionally stable.

3.2.2 Alternating Direction Implicit Method

In explicit methods the unknown states at a later time can be found based on the known current states as shown by Eq. (3.7). For implicit methods, the states of a system at a later time are defined by a set of coupled equations. So, either a matrix inversion or iterative technique is required to obtain the states of a system. The main advantage of implicit methods is that it allows a larger time step, which in turns can lead to a faster without compromising accuracy [80]. The matrix inversion involved in the implicit method is computationally intensive because it is involved with the banded matrix that requires substantial amount of computer time. The alternating directional implicit (ADI) method is an implicit method for parabolic and elliptic partial differential equations. The major advantage of the ADI method compared with other implicit methods is that since it involves only tri-diagonal systems, the matrix inversion can be achieved efficiently using Thomas' algorithm [70]. For many dynamic

problems the associated Fokker-Planck equation can be written in the following form

$$\frac{\partial p}{\partial t} = \sum_{n=1}^{M} \left[a_n \frac{\partial p}{\partial \alpha_n} + b_n \frac{\partial p}{\partial \dot{\alpha}_n} + \frac{Q_n}{2} \frac{\partial^2 p}{\partial \dot{\alpha}_n^2} \right], \qquad n = 1, 2, \ldots, N/2 \tag{3.10}$$

where a_n and b_n could be constants and/or polynomial of α and/or $\dot{\alpha}$, N is the size of the state vector. The operators in the ADI method are:

$$A_i = a_i \frac{\partial}{\partial \alpha_i}, \qquad i = 1, 2, \ldots, N/2 \tag{3.11}$$

and

$$A_j = b_j \frac{\partial}{\partial \dot{\alpha}_j} + \frac{Q_j}{2} \frac{\partial^2}{\partial \dot{\alpha}_j^2}, \qquad j = i = 1, 2, \ldots, N/2 \tag{3.12}$$

With these operators, equation (3.10) can be rewritten as

$$\frac{\partial p}{\partial t} = \sum_{k=1}^{N} A_k p \tag{3.13}$$

where the index for the operators is $k = 1, \ldots, N$. After discretizing Eq. (3.13) in time but not in space using the implicit Euler scheme the equation becomes as

$$\frac{p^{t+1} - p^t}{\Delta t} = \sum_{k=1}^{N} A_k p^{t+1} \tag{3.14}$$

Rearranging the terms in Eq. (3.14) leads to

$$\left(\mathbf{I} - \Delta t \sum_{k=1}^{N} A_k \right) p^{t+1} = p^t \tag{3.15}$$

By inverting $\left(\mathbf{I} - \Delta t \sum_{k=1}^{N} A_k \right)$, p^{t+1} can be obtained, but the direct inversion of $\left(\mathbf{I} - \Delta t \sum_{k=1}^{N} A_k \right)$ can be computationally expensive. This inversion process can be

simplified by using the following identity [70]

$$\prod_{k=1}^{N} (\mathbf{I} \pm \Delta t A_k) \cong \mathbf{I} \pm \Delta t \sum_{k=1}^{N} A_k + \Delta t^2 \sum_{k<l} A_k A_l \tag{3.16}$$

By neglecting the high order term, $\Delta t^2 \sum_{k<l} A_k A_l$, Eq. (3.15) can be written as

$$\prod_{k=1}^{N} (\mathbf{I} - \Delta t A_k)\, p^{t+1} = p^t \quad \text{or} \quad p^{t+1} = \prod_{k=1}^{N} (\mathbf{I} - \Delta t A_k)^{-1}\, p^t \tag{3.17}$$

So, the ADI scheme for a general N-order system is

$$p^{t+\frac{k}{N}} = (\mathbf{I} - \Delta t A_k)^{-1}\, p^{t+\frac{k-1}{N}} \qquad k = 1 \ldots N \tag{3.18}$$

The discretization in space is done with an up-wind differencing for the first order spatial derivatives and the central differencing scheme for the second order terms

$$A_k p^{t+\frac{k-1}{N}} = a_i \frac{\partial p^{t+\frac{k-1}{N}}}{\partial \alpha_i} = \frac{a_i}{\Delta \alpha_i} \begin{cases} \left(p_{i+1}^{t+\frac{k-1}{N}} - p_i^{t+\frac{k-1}{N}} \right) & a_i \geq 0 \\ \left(p_i^{t+\frac{k-1}{N}} - p_{i-1}^{t+\frac{k-1}{N}} \right) & a_i < 0 \end{cases} \tag{3.19}$$

$$A_k p^{t+\frac{k-1}{N}} = a_i \frac{\partial p^{t+\frac{k-1}{N}}}{\partial \alpha_i} = \frac{a_i}{\Delta \alpha_i} \begin{cases} \left(p_{i+1}^{t+\frac{k-1}{N}} - p_i^{t+\frac{k-1}{N}} \right) & a_i \geq 0 \\ \left(p_i^{t+\frac{k-1}{N}} - p_{i-1}^{t+\frac{k-1}{N}} \right) & a_i < 0 \end{cases} \tag{3.20}$$

After discretizing in space $(\mathbf{I} - \Delta t A_k)$ becomes a tridiagonal matrix and can be inverted very quickly using Thomas' algorithm [70]. The operators in the form of

Eq. (3.19) with boundary conditions of p_{i_1} and $p_{i_{max}}$ can be rewritten as when $a_i \geq 0$

$$\begin{bmatrix} p_{i_1}^{t+\frac{k}{N}} \\ p_{i_2}^{t+\frac{k}{N}} \\ \vdots \\ \vdots \\ p_{i_{\max}-2}^{t+\frac{k}{N}} \\ p_{i_{\max}-1}^{t+\frac{k}{N}} \end{bmatrix} = \begin{bmatrix} M & N & 0 & \cdots & \cdots & 0 \\ 0 & M & N & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & N & 0 \\ 0 & \cdots & 0 & 0 & M & N \\ 0 & \cdots & 0 & 0 & 0 & M \end{bmatrix}^{-1} \begin{bmatrix} p_{i_1}^{t+\frac{k-1}{N}} \\ p_{i_2}^{t+\frac{k-1}{N}} \\ \vdots \\ \vdots \\ p_{i_{\max}-2}^{t+\frac{k-1}{N}} \\ p_{i_{\max}-1}^{t+\frac{k-1}{N}} + a \cdot \frac{\Delta t}{\Delta \alpha} p_{i_{\max}}^{t+\frac{k-1}{N}} \end{bmatrix} \tag{3.21}$$

where $M = 1 + a_i \dfrac{\Delta t}{\Delta \alpha_i}$ and $N = -a_i \dfrac{\Delta t}{\Delta \alpha_i}$.

When $a_i < 0$

$$\begin{bmatrix} p_{i_1}^{t+\frac{k}{N}} \\ p_{i_2}^{t+\frac{k}{N}} \\ \vdots \\ \vdots \\ p_{i_{\max}-2}^{t+\frac{k}{N}} \\ p_{i_{\max}-1}^{t+\frac{k}{N}} \end{bmatrix} = \begin{bmatrix} M & 0 & 0 & \cdots & \cdots & 0 \\ L & M & 0 & 0 & \cdots & 0 \\ 0 & L & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & L & M & 0 \\ 0 & \cdots & 0 & 0 & L & M \end{bmatrix}^{-1} \begin{bmatrix} p_{i_1}^{t+\frac{k-1}{N}} - a \cdot \frac{\Delta t}{\Delta \alpha} p_{i_0}^{t+\frac{k-1}{N}} \\ p_{i_2}^{t+\frac{k-1}{N}} \\ \vdots \\ \vdots \\ p_{i_{\max}-2}^{t+\frac{k-1}{N}} \\ p_{i_{\max}-1}^{t+\frac{k-1}{N}} \end{bmatrix} \tag{3.22}$$

where $L = a_i \dfrac{\Delta t}{\Delta \alpha_i}$ and $M = 1 - a_i \dfrac{\Delta t}{\Delta \alpha_i}$. When $a_i \geq 0$ and $a_i < 0$ the operators in the form of Eq. (3.20) with boundary conditions of p_{j_1} and $p_{j_{max}}$ can be rewritten as

Eq. (3.23) and Eq. (3.24) accordingly

$$\begin{bmatrix} p_{j_1}^{t+\frac{k+1}{N}} \\ p_{j_2}^{t+\frac{k+1}{N}} \\ \vdots \\ \vdots \\ p_{j_{\max}-2}^{t+\frac{k+1}{N}} \\ p_{j_{\max}-1}^{t+\frac{k+1}{N}} \end{bmatrix} = \begin{bmatrix} M & L & 0 & \cdots & \cdots & 0 \\ -L & M & N & 0 & \cdots & 0 \\ 0 & -L & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & N & 0 \\ 0 & \cdots & 0 & -L & M & N \\ 0 & \cdots & 0 & 0 & -L & M \end{bmatrix}^{-1} \begin{bmatrix} p_{j_1}^{t+\frac{k}{N}} + a \cdot p_{j_0}^{t+\frac{k}{N}} \\ p_{j_2}^{t+\frac{d-1}{N}} \\ \vdots \\ \vdots \\ p_{j_{\max}-2}^{t+\frac{k}{N}} \\ p_{j_{\max}-1}^{t+\frac{k}{N}} + b \cdot p_{j_{\max}}^{t+\frac{k}{N}} \end{bmatrix} \tag{3.23}$$

where $L = \dfrac{Q_j \Delta t}{2\Delta\dot{\alpha}_j^2}$, $M = 1 + b_j \dfrac{\Delta t}{\Delta\dot{\alpha}_j} + \dfrac{Q_j \Delta t}{\Delta\dot{\alpha}_j^2}$ and $N = -\left(b_j \dfrac{\Delta t}{\Delta\dot{\alpha}_j} + \dfrac{Q_j \Delta t}{2\Delta\dot{\alpha}_j^2} \right)$

$$\begin{bmatrix} p_{j_1}^{t+\frac{k+1}{N}} \\ p_{j_2}^{t+\frac{k+1}{N}} \\ \vdots \\ \vdots \\ p_{j_{\max}-2}^{t+\frac{k+1}{N}} \\ p_{j_{\max}-1}^{t+\frac{k+1}{N}} \end{bmatrix} = \begin{bmatrix} M & -N & 0 & \cdots & \cdots & 0 \\ L & M & -N & 0 & \cdots & 0 \\ 0 & L & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -N & 0 \\ 0 & \cdots & 0 & L & M & -N \\ 0 & \cdots & 0 & 0 & L & M \end{bmatrix}^{-1} \begin{bmatrix} p_{j_1}^{t+\frac{k}{N}} - a \cdot p_{j_0}^{t+\frac{k}{N}} \\ p_{j_2}^{t+\frac{k}{N}} \\ \vdots \\ \vdots \\ p_{j_{\max}-2}^{t+\frac{k}{N}} \\ p_{j_{\max}-1}^{t+\frac{k}{N}} + c \cdot p_{j_{\max}}^{t+\frac{k}{N}} \end{bmatrix} \tag{3.24}$$

where $L = b_j \dfrac{\Delta t}{\Delta\dot{\alpha}_j} - \dfrac{Q_j \Delta t}{\Delta\dot{\alpha}_j^2}$, $M = 1 - b_j \dfrac{\Delta t}{\Delta\dot{\alpha}_j} + \dfrac{Q_j \Delta t}{\Delta\dot{\alpha}_j^2}$ and $N = \dfrac{Q_j \Delta t}{2\Delta\dot{\alpha}_j^2}$. As mentioned before, all the matrix inversions in Eq. (3.21) through Eq. (3.24) involve only tri-diagonal matrices.

3.2.3 Moving Domain

Solving the FPE numerically could be computationally prohibitive even for a low dimension system due to the fact that a PDF is a function defined over an infinite domain even with the fast numerical method such as the ADI. A large domain is used to make sure that the majority of the PDF is covered throughout numerical

propagation [15]. However, this approach will cause the numerical solver to spend substantial time in the part of domain that carries not much useful information to the solution of the FPE. Consequently, it is essential to truncate the size of the domain properly without sacrificing the accuracy. The size of the truncated domain should be large enough to contain the entire PDF and sufficiently small so that the computational effort would not be wasted on the part of the domain that dose not contribute much [15] as illustrated as in the Fig. 3.3. As the PDF evolves along with

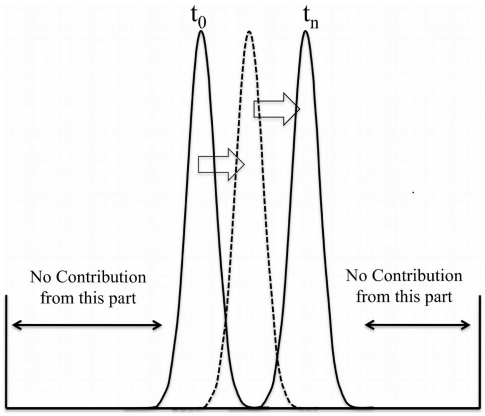


Figure 3.3: The Size of the Domain is too Big

the system dynamics, the optimal domain for the current time can lose its optimality as PDF evolves to the next time step. The figure 3.4 provides the illustration of this problem. In the figure, the domain is optimal for the PDF at t_0 but it clearly loses its optimality at t_n as most of its right side of the PDF drifted to the outside of the domain. Challa and Bar-Shalom [15] developed a moving domain scheme that adaptively calculates the size and the location of the domain in a recursive manner using Chebyshev's inequality combined with the moment evolution equation [15]. This method predicts the future domain at each measurement update. The union of the current domain and the predicted domain becomes the domain for the next propagation. The advantage of Challa's method comes from the Chebyshev's inequality theory which allows the size of the domain to be adaptively truncated at each measurement update. In Musick's work [48, 49], after each measurement

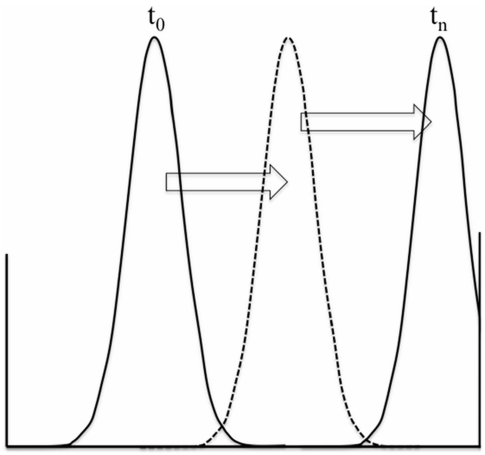
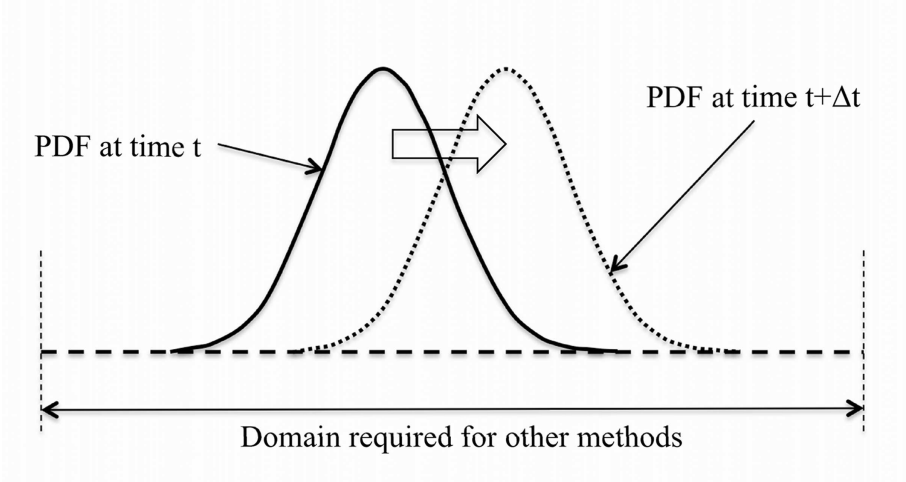
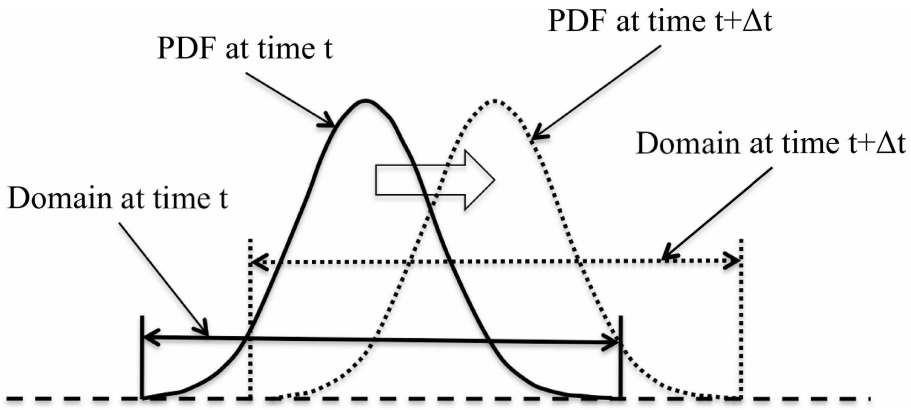


Figure 3.4: The drift of the PDF from time t_0 to $t_n$time

the domain with a fixed size is adjusted based on the estimation so that the PDF is located at the near center of the domain [48, 49]. However if the size of the PDF becomes much smaller than it is at the beginning, the computational effort will be wasted on the part of domain with little contribution as shown in the Fig. 3.3. The common aspect of both methods is that they do not adjust the domain during the evaluation of the FPE. The adjustment happens only after each measurement update. If the domain can be adjusted continuously during the numerical evaluation of the FPE based on the movement of the PDF, the size of the domain can be even smaller. This method of continuously adjusting the domain is developed and used here to reduce the computational load. Figure 3.5 illustrates the basic idea of the domain adjustment scheme used in this dissertation. Note that the domain size in the Fig. 3.5(b) is smaller than the domain used in other methods as shown in the Fig. 3.5(a). In Fig. 3.5(a), the domain has to cover the whole possible PDF range between measurement updates, whereas in the Fig. 3.5(b), the domain only needs to cover the possible PDF range between each time step instead of the measurement update. Tracking the movement of the PDF is achieved by taking advantage of the fact that the PDF studied here is not a multimodal distribution. Since the state PDF has only one peak, the movement of the PDF was followed by comparing the



(a) The drift of the PDF from the center of the domain in Δt and the domain needs to cover the whole range between the measurement updates



(b) The drift of the PDF from the center of the domain in Δt and the domain is adjusted during the FPE propagation between measurement updates. The necessary domain size is smaller than one in Fig.3.5(a)

Figure 3.5: Drift of PDF

location of the peak of the PDF and the center of the domain as it is in Fig. 3.5(a). In Fig. 3.5, Δt is the time step size for the FPE propagation. If the distance between the center of the domain and the peak of the PDF is larger than a pre-defined value, during the numerical evaluation along the time, the domain is adjusted such that the peak of the PDF is located at the desired location on the domain, i.e., the center of the domain (Fig.3.5(b)). As compared with the methods in Challa and Bar-Shalom [15] and Musick et al. [48, 49], since the domain can be adjusted during the FPE propagation, the domain size is much smaller than the case where the adjustment is

only taken after each measurement update. To illustrate the effect of the proposed moving domain scheme, let's suppose the domain, in each step of the propagation between two measurements, needs to have 10 grids in each direction to adequately accommodate the PDF. Then, the total numbers of evaluations needed for 2, 4 and 6 dimensional system are 10^2, 10^4 and 10^6 iterations correspondingly. Suppose the domain in each step is only decreased by 10% (or by one grid) through translating the domain during the evaluation of the FPE, the total numbers of iterations needed for 2, 4 and 6 dimensional system are 81, 6561 and 531441 iterations, which are 19%, 34% and 47% less than the original domain.

3.2.4 Measurement Update for Finite Difference Filter

If the measurements are conditioned only on the state of the system, Bayes' formula is used to update the conditional probability density function $p(\mathbf{x}, t_k|\mathbf{Y}_{k-1})$ which is the solution of the FPE to obtain $p(\mathbf{x}, t_k|\mathbf{Y}_k)$.

$$p(\mathbf{x}, t_k|\mathbf{Y}_{t_k}) = \frac{p(\mathbf{x}, t_k|\mathbf{Y}_{t_{k-1}})p(\mathbf{y}_k|\mathbf{x})}{\int p(\xi|\mathbf{Y}_{t_{k-1}})p(\mathbf{y}_t|\xi)d\xi} \tag{3.25}$$

Equations (3.6) and (3.25) represent the predictor and corrector equations for the exact optimal nonlinear filtering. The PDF $p(\mathbf{y}_k|\mathbf{x})$ is defined by the characteristics of the sensor and is usually assumed to have a Gaussian distribution as

$$p(\mathbf{y}_k|\mathbf{x}) = \frac{1}{(2\pi)^{m/2}|\mathbf{R}|^{1/2}} e^{\{-\frac{1}{2}[\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k)]^T \mathbf{R}^{-1}[\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k)]\}} \tag{3.26}$$

Now the estimation of the each state can be made by calculating the following integral

$$\hat{x}_i(t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i p(\mathbf{x}, t_k|\mathbf{Y}_k) \prod_{j=1}^{n} dx_i \qquad i = 1, 2, \ldots, n \tag{3.27}$$

This produces the optimal estimate in the sense of MMSE that minimizes the mean square error given by

$$J_i = \frac{1}{2}E[(\hat{x}_i - x_i)^2] \qquad i = 1, 2, \ldots, n \tag{3.28}$$

3.3 Filter based on the Direct Quadrature Method Of Moments

3.3.1 Direct Quadrature Moments Of Method

The DQMOM method, originally investigated by Marchisio and Fox for the population balance problem [57] and further developed for solution of the Fokker-Planck equation [6, 7], is illustrated in terms of nonlinear filtering [75]. First, let us define the state conditional PDF as a summation of a multi-dimensional Dirac delta function

$$p(\mathbf{x}_t|\mathbf{Y}_k) = \sum_{\alpha=1}^{N} w_\alpha \prod_{j=1}^{N_s} \delta[x_j - \langle x_j \rangle_\alpha] \tag{3.29}$$

where w_α, $\alpha = 1, ..., N$ is the corresponding weight for node α, and $\langle x_j \rangle_\alpha$, $j = 1, ..., Ns$ is the property vector of node α called "abscissas" N and N_s are the number of nodes used in the PDF representation and the size of the state vector. In this representation, there are total $N(N_s + 1)$ unknown variables which will be solved through moment constraints. The weights and abscissas will be computed by substituting Eq. (3.29) into Eq. (3.6), then the left hand side (LHS) of Eq. (3.6) becomes

$$\begin{aligned}
\frac{\partial p}{\partial t} &= \frac{\partial}{\partial t}\{\sum_{\alpha=1}^{N} w_\alpha \prod_{j=1}^{N_s} \delta[x_j - \langle x_j \rangle_\alpha]\} \\
&= \sum_{\alpha=1}^{N} \left(\frac{\partial w_\alpha}{\partial t} \right) \prod_{j=1}^{N_s} \delta[x_j - \langle x_j \rangle_\alpha] - \sum_{\alpha=1}^{N} w_\alpha \sum_{j=1}^{N_s} \prod_{k=1,k\neq j}^{N_s} \delta[x_j - \langle x_j \rangle_\alpha] \frac{\partial \delta_{j\alpha}}{\partial \langle x_j \rangle_\alpha} \frac{\partial \langle x_j \rangle_\alpha}{\partial t} \\
&= \sum_{\alpha=1}^{N} \prod_{j=1}^{N_s} \delta_{j\alpha} \left(\frac{\partial w_\alpha}{\partial t} \right) - \sum_{\alpha=1}^{N} \sum_{j=1}^{N_s} \prod_{k=1,k\neq j}^{N_s} w_\alpha \delta_{k\alpha} \delta'_{j\alpha} \frac{\partial \langle x_j \rangle_\alpha}{\partial t}
\end{aligned} \tag{3.30}$$

where $\delta_{j\alpha} = \delta[x_j - \langle x_j \rangle_\alpha]$ and $\delta'_{j\alpha} = \partial \delta_{j\alpha} / \partial \langle x_j \rangle_\alpha$. If the weighted abscissas $\zeta_{j\alpha} = w_\alpha \langle x_j \rangle_\alpha$ is introduced, after some manipulations, Eq. (3.30) can be rewritten as

$$\frac{\partial p}{\partial t} = \sum_{\alpha}^{N} \left[\prod_{j=1}^{Ns} \delta_{j\alpha} \left(\frac{\partial w_\alpha}{\partial t} \right) + \sum_{j=1}^{N} \prod_{k=1, k\neq j}^{N_s} \langle x_j \rangle_\alpha \delta_{k\alpha} \delta'_{j\alpha} \frac{\partial w_\alpha}{\partial t} \right] - \sum_{\alpha=1}^{N} \sum_{j=1}^{N_s} \prod_{k=1, k\neq j}^{N_s} \delta_{k\alpha} \delta'_{j\alpha} \frac{\partial \zeta_{j\alpha}}{\partial t} \tag{3.31}$$

Notice that w_α, $\zeta_{j\alpha}$, and $\delta_{j\alpha}$ are functions of only time, thus

$$\frac{dp}{dt} = \sum_{\alpha}^{N} \left[\prod_{j=1}^{Ns} \delta_{j\alpha} \left(\frac{dw_\alpha}{dt} \right) + \sum_{j=1}^{N} \prod_{k=1, k\neq j}^{N_s} \langle x_j \rangle_\alpha \delta_{k\alpha} \delta'_{j\alpha} \frac{dw_\alpha}{dt} \right] - \sum_{\alpha=1}^{N} \sum_{j=1}^{N_s} \prod_{k=1, k\neq j}^{N_s} \delta_{k\alpha} \delta'_{j\alpha} \frac{d\zeta_{j\alpha}}{dt} \tag{3.32}$$

Let us define

$$dw_\alpha/dt = a_\alpha,\ \alpha = 1, \ldots, N \text{ and } d\zeta_{j\alpha}/dt = b_{j\alpha},\ j = 1, \ldots, N_s; \alpha = 1, \ldots, N \tag{3.33}$$

Eq. (3.32) can be further simplified as

$$\frac{dp}{dt} = \sum_{\alpha}^{N} \left[\prod_{j=1}^{Ns} \delta_{j\alpha} + \sum_{j=1}^{N} \prod_{k=1, k\neq j}^{N_s} \langle x_j \rangle_\alpha \delta_{k\alpha} \delta'_{j\alpha} \right] a_\alpha - \sum_{\alpha=1}^{N} \left[\sum_{j=1}^{N_s} \prod_{k=1, k\neq j}^{N_s} \delta_{k\alpha} \delta'_{j\alpha} \right] b_{j\alpha} \tag{3.34}$$

Now let the right hand side(RHS) of the FPE (Eq. (3.6)) defined to be

$$\mathbf{S_x}(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial [p \mathbf{f}_i]}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 [p(\mathbf{GQG^T})_i j]}{\partial x_i \partial x_j} \tag{3.35}$$

The Eq. (3.35) can be written in terms of the multi-variable Dirac delta function as

$$\mathbf{S_x}(\mathbf{x}) = \sum_{\alpha=1}^{N} \left(\prod_{j=1}^{N_s} \delta_{j\alpha} \right) a_\alpha + \sum_{\alpha=1}^{N} \sum_{j=1}^{N_s} \prod_{k=1, k\neq j}^{N_s} \langle x_j \rangle_\alpha \delta_{k\alpha} \delta'_{j\alpha} a_\alpha - \sum_{\alpha=1}^{N} \left[\sum_{j=1}^{N_s} \prod_{k=1, k\neg j}^{N_s} \delta_{k\alpha} \delta'_{j\alpha} \right] b_{j\alpha} \tag{3.36}$$

There are total $N(1+N_s)$ parameters in Eq. (3.36) need to be found to construct the conditional PDE $p(\mathbf{x}_t|\mathbf{Y}_k)$: a_α, $\alpha = 1, \ldots, N$ and $b_{j\alpha}$, $j = 1, \ldots, N_s$, $\alpha = 1, \ldots, N$. In general, DQMOM method applies an independent set of moments that user wish

to control to construct $N(1+N_s)$ a differential algebraic equations (DAEs). Given the following three Dirac delta function properties

$$\int_{-\infty}^{+\infty} x_k \delta(x - \langle x \rangle_\alpha) dx = \langle x \rangle_\alpha^k \tag{3.37}$$

$$\int_{-\infty}^{+\infty} x_k \delta'(x - \langle x \rangle_\alpha) dx = -k \langle x \rangle_\alpha^{k-1} \tag{3.38}$$

and

$$\int_{-\infty}^{+\infty} x_k \delta''(x - \langle x \rangle_\alpha) dx = k(k-1) \langle x \rangle_\alpha^{k-2} \tag{3.39}$$

The $k_1, k_2, \ldots, k_{N_s}$ moment of the Eq. (3.36) can be derived as followed

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} x_1^{k_1} \cdots x_{N_s}^{N_s} \left[\sum_{\alpha}^{N} \prod_{j=1}^{Ns} \delta_{j\alpha} a_\alpha \right] \prod_{I=1}^{Ns} dx_I \\
&+ \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} x_1^{k_1} \cdots x_{N_s}^{N_s} \left[\sum_{\alpha}^{N} \sum_{j=1}^{N_s} \prod_{k=1 k \neg j}^{Ns} \langle x_j \rangle_\alpha \delta_{k\alpha} \delta'_{j\alpha} a_\alpha \right] \prod_{I=1}^{Ns} dx_I \\
&- \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} x_1^{k_1} \cdots x_{N_s}^{N_s} \left[\sum_{\alpha}^{N} \left[\sum_{j=1}^{N_s} \prod_{k=1 k \neg j}^{Ns} \delta_{k\alpha} \delta'_{j\alpha} \right] b_{j\alpha} \right] \prod_{I=1}^{Ns} dx_I \\
&= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} x_1^{k_1} \cdots x_{N_s}^{N_s} \left[\mathbf{S}_{\mathbf{x}}(\mathbf{x}) \right] \prod_{I=1}^{Ns} dx_I
\end{aligned} \tag{3.40}$$

The first term in the LHS of Eq. (3.40)

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} x_1^{k_1} \cdots x_{N_s}^{N_s} \left[\sum_{\alpha}^{N} \prod_{j=1}^{Ns} \delta_{j\alpha} a_\alpha \right] \prod_{I=1}^{Ns} dx_I \\
&= \sum_{\alpha=1}^{N} a_\alpha \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} x_1^{k_1} \cdots x_{N_s}^{N_s} (\delta_{1\alpha} \delta_{2\alpha} \cdots \delta_{N_s \alpha}) dx_1 \cdots dx_{N_s} \\
&= \sum_{\alpha=1}^{N} \left[\prod_{j=1}^{N_s} \langle x_j \rangle \right] a_\alpha
\end{aligned} \tag{3.41}$$

where as the second term in the LHS of Eq. (3.40) is derived to be

$$\begin{aligned}
&\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} x_1^{k_1}\cdots x_{N_s}^{N_s}\left[\sum_{\alpha}^{N}\sum_{j=1}^{N_s}\prod_{k=1 k\neg j}^{Ns}\langle x_j\rangle_\alpha \delta_{k\alpha}\delta'_{j\alpha}a_\alpha\right]\prod_{I=1}^{Ns}dx_I \\
&=\sum_{\alpha=1}^{N}a_\alpha\sum_{j=1}^{N_s}\int_{-\infty}^{+\infty}\left\{\prod_{m=1,m\neq j}^{N_s}x_m^{k_m}\prod_{k=1,k\neq j}^{N_s}\delta_{k\alpha}\left[\int_{-\infty}^{+\infty}x_j^{k_j}\langle x_j\rangle_\alpha\delta'_{j\alpha}dx_j\right]\prod_{I=1,I\neq j}^{N_s}dx_I\right\} \\
&=\sum_{\alpha=1}^{N}a_\alpha\sum_{j=1}^{N_s}\left[(-k_j)\langle x_j\rangle_\alpha^{k_j}\right]\int_{-\infty}^{+\infty}\left\{\prod_{m=1,m\neq j}^{N_s}x_m^{k_m}\prod_{k=1,k\neq j}^{N_s}\delta_{k\alpha}\prod_{I=1,I\neq j}^{N_s}dx_I\right\} \qquad (3.42)\\
&=-\sum_{\alpha=1}^{N}\left(\sum_{j=1}^{N_s}k_j\prod_{k=1}^{N_s}\langle x_k\rangle_\alpha^{k_k}\right)a_\alpha
\end{aligned}$$

In the same way, the third term of the LHS of Eq. (3.40) can be written as

$$-\sum_{\alpha=1}^{N}\sum_{j=1}^{N_s}b_{j\alpha}\int_{-\infty}^{+\infty}x_1^{k_1}\cdots x_{N_s}^{k_{N_s}}\delta'_{j\alpha}\left(\prod_{k=1,k\neq j}^{N_s}\delta_{k\alpha}\right)dx_1\ldots dx_{N_s} \qquad (3.43)$$

$$=\sum_{\alpha=1}^{N}\sum_{j=1}^{N_s}k_j\langle x_j\rangle_\alpha^{k-1}\prod_{k=1,k\neq j}^{N_s}\langle x_j\rangle_\alpha^{k}b_{j\alpha} \qquad (3.44)$$

The $k_1,\ldots,k_{N_s}$ moment of the RHS of Eq. (3.40) is derived to be

$$\begin{aligned}
\overline{\mathbf{S}}_{k_1,\ldots,k_{N_s}} &= \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}x_1^{k_1}\cdots x_{N_s}^{k_{N_s}}\left[\mathbf{S}_{\mathbf{x}}(\mathbf{x})\right]dx_1\ldots dx_{Ns} \\
&= -\sum_{i=1}^{N_s}\int_{-\infty}^{+\infty}x_1^{k_1}\cdots x_{N_s}^{N_s}\left[\frac{\partial p\mathbf{f}_i}{\partial x_i}\right]dx_1\ldots dx_{Ns} \qquad (3.45)\\
&+ \int_{-\infty}^{+\infty}x_1^{k_1}\cdots x_{N_s}^{k_{N_s}}\left[\frac{1}{2}\sum_{i=1}^{N_s}\sum_{j=1}^{N_s}\frac{\partial^2[p(\mathbf{GQG^T})_{ij}]}{\partial x_i\partial x_j}\right]dx_1\ldots dx_{Ns} \\
&= \overline{\mathbf{S}}^1_{k_1,\ldots,k_{N_s}}+\overline{\mathbf{S}}^2_{k_1,\ldots,k_{N_s}}
\end{aligned}$$

where

$$\begin{aligned}
\overline{\mathbf{S}}^{1}_{k_1,\ldots,k_{N_s}} &= -\sum_{i=1}^{N_s}\int_{-\infty}^{+\infty} x_1^{k_1}\cdots x_{N_s}^{N_s}\left[\frac{\partial p\mathbf{f}_i}{\partial x_i}\right]dx_1\ldots dx_{Ns} \\
&= -\sum_{i=1}^{N_s}\int_{-\infty}^{+\infty} x_1^{k_1}\cdots x_{N_s}^{N_s}\frac{\partial}{\partial x_i}\left[\mathbf{f}_i(\mathbf{x})\sum_{\alpha}^{N} w_\alpha(t)\prod_{j=1}^{N_s}\delta_{j\alpha}\right]dx_1\ldots dx_{Ns} \qquad (3.46)\\
&= \sum_{i=1}^{N_s}\sum_{\alpha=1}^{N} k_i w_\alpha(t)\langle x_1\rangle_\alpha^{k_1}\cdots\langle x_{i-1}\rangle_\alpha^{k_{i-1}}\langle x_i\rangle_\alpha^{k_i}\langle x_{i+1}\rangle_\alpha^{k_{i+1}}\cdots\langle x_{N_s}\rangle_\alpha^{k_{N_s}}\mathbf{f}_i(\langle x_1\rangle_\alpha,\ldots\langle x_{N_s}\rangle_\alpha
\end{aligned}$$

When $i \neq j$, $\overline{\mathbf{S}}^{2}_{k_1,\ldots,k_{N_s}}$ is derived as

$$\begin{aligned}
\overline{S}^{2}_{k_1,\ldots,k_{N_s}} &= \int_{-\infty}^{+\infty} x_1^{k_1}\cdots x_{N_s}^{k_{N_s}}\left[\frac{1}{2}\sum_{i=1}^{N_s}\sum_{j=1}^{N_s}\frac{\partial^2[p(\mathbf{GQG^T})_{ij}]}{\partial x_i\partial x_j}\right]dx_1\ldots dx_{Ns} \\
&= \sum_{i=1}^{N_s}\sum_{j=1}^{N_s}\int_{-\infty}^{+\infty} x_1^{k_1}\cdots x_{N_s}^{k_{N_s}}\left[\frac{\partial^2 p[D(\mathbf{x})]_{ij}}{\partial x_i\partial x_j}\right] \qquad (3.47)\\
&= \sum_{i=1}^{N_s}\sum_{j=1}^{N_s}\sum_{\alpha=1}^{N} w_\alpha k_i k_j\left[\prod_{k=1}^{N_s}\langle x_k\rangle_\alpha^{k_k}\right]/\langle x_i\rangle_\alpha\langle x_j\rangle_\alpha[D(\mathbf{x})]_{ij}|_{\langle x_1\rangle_\alpha,\ldots,\langle x_{N_s}\rangle_\alpha}
\end{aligned}$$

where as when $i = j$,$\overline{S}^{2}_{k_1,\ldots,k_{N_s}}$ is derived as,

$$\begin{aligned}
\overline{S}^{2}_{k_1,\ldots,k_{N_s}} &= \int_{-\infty}^{+\infty} x_1^{k_1}\cdots x_{N_s}^{k_{N_s}}\frac{\partial^2 p[D(\mathbf{x})]_{ii}}{\partial x_i^2}dx_1\ldots dx_{Ns} \qquad (3.48)\\
&= \sum_{\alpha=1}^{N} w_\alpha k_i(k_i-1)\left[\prod_{k=1}^{N_s}\langle x_k\rangle_\alpha^{k_k}\right]/\langle x_i\rangle_\alpha^2[D(\mathbf{x})]_{ij}|_{\langle x_1\rangle_\alpha,\ldots,\langle x_{N_s}\rangle_\alpha}
\end{aligned}$$

Notice that $D(\mathbf{x}) = (1/2)GQG^T$. Thus, the $N(1+N_s)$ DAEs can be constructed using a set of independent moments constraints $k_1,\ldots,k_{N_s}$ as

$$\overline{\mathbf{S}}_{k_1,\ldots,k_{N_s}} = \sum_{\alpha=1}^{N}\left[\left(1-\sum_{j=1}^{N_s}k_j\right)\prod_{k=1}^{N_s}\langle x_k\rangle_\alpha^{k_k}\right]a_\alpha + \sum_{\alpha=1}^{N}\sum_{j=1}^{N_s}k_j\langle x_j\rangle_\alpha^{k_j-1}\prod_{k=1,k\neq j}^{N_s}\langle x_k\rangle_\alpha^{k_k}b_{j\alpha} \qquad (3.49)$$

For example, if the number of states is N_s=2 and the number of nodes used in the multi-dimensional Dirac delta function is N=2, there will be $N(1+N_s)$=6 unknown parameters in Eq. (3.49). In order solve these six DAEs, the following six moments

constraints, $(k_1, k_2) = (0,0), (1,0), (0,1), (2,0), (1,1), (0,2)$ can be applied such that there are enough equations for solving $a_\alpha,\ \alpha = 1,2$ and $b_{j\alpha},\ j = 1,2;\ \alpha = 1,2$ explicitly. Typically, the precision of the estimation and the computational cost will be higher when the number of nodes increases. The selected moment constraint $k_1, k_2, \ldots, k_{N_s}$ will guarantee the PDF approximated by the Eq. (3.29) has exact value for this moment of the PDF. For typical estimation problem, the accuracy of the first moment (e.g. minimum mean-square estimate (MMSE) estimates of any state variables or functions of state variables $\phi(\mathbf{x})$ can be obtained) is automatically guaranteed. For simplicity, Eq. (3.49) can be rewritten in a matrix form as

$$\mathbf{A}\boldsymbol{\mu} = \mathbf{s} \tag{3.50}$$

where the unknown parameters are

$$\boldsymbol{\mu} = [a_1, a_2, \ldots, a_n, b_{11}, b_{12}, \ldots, b_{1N}, \ldots, b_{N_s 1}, b_{N_s 2}, \ldots, b_{N_s N}]^T \in \mathbb{R}^{N(1+N_s)\times 1} \tag{3.51}$$

and matrix A can be derived from Eq. (3.49) as a nonlinear function of the abscissas. The moment constraints are

$$\mathbf{s} = [\overline{S}_{0,\ldots,0}, \overline{S}_{1,\ldots,0}, \ldots] \in \mathbb{R}^{N(1+N_s)\times 1} \tag{3.52}$$

The FPE, apartial differential equation, is reduced to a differential algebraic equation which is much easier to calculate.

3.3.2 Update Schemes for the DQMOM based Nonlinear Filters

Update based on Bayes' Formula

Once the weights and abscissas in the predictor PDF are found through the DQMOM propagation, the updated conditional PDF can be found using the new measurement

y_{k+1} made at the time instant $k+1$. Substitute Eq. (3.29) into Eq. (3.25), the DQMOM based Bayes' equation can be derived as following: (1) the updated weights are

$$w_\alpha(t, Y_{k+1}) = \frac{w_\alpha(t, Y_k) p\left(y_{k+1} | \langle x_1 \rangle_\alpha, ..., \langle x_{N_s} \rangle_\alpha\right)}{\sum_{\alpha=1}^{N} w_\alpha(t, Y_k) p\left(y_{k+1} | \langle x_1 \rangle_\alpha, ..., \langle x_{N_s} \rangle_\alpha\right)}, \ \alpha = 1, ..., N \tag{3.53}$$

and (2) the abscissas are unchanged as

$$\left\langle x_j \right\rangle_\alpha (t, Y_{k+1}) = \left\langle x_j \right\rangle_\alpha (t, Y_k)), \ \alpha = 1, ..., N, \ j = 1, ..., N_s \tag{3.54}$$

With this update method the "degeneracy" phenomenon, similar to the one exists in a typical particle filter, appears because in this algorithm only the weight is updated and the abscissas remain the same [76]. The un-updated abscissas might be propagated into the tail of the PDF where no significant statistical meaning is carried. Thus the effect of measurement update is limited and the filter struggles to correct errors that are too large to be fixed by the weight update only. In next two sections, the update mechanisms employed to mitigate the "degeneracy" phenomenon.

Update through the Extended Kalman Filter

The prediction of the states $\hat{\mathbf{x}}^- = [\hat{x}_i^-]_{i=1,...,N_s}$ and the error covariance matrix $\mathbf{P}^- = [P_{ij}^-]_{i=1,...,N_s, j=1,...,N_s}$ at time step t_k are calculated from the abscissas and weights propagated through the DQMOM as

$$\hat{x}_i^- = \sum_{\alpha=1}^{N} w_\alpha \langle x_i \rangle_\alpha \qquad i = 1, \ldots, N_s \tag{3.55}$$

and

$$P_{ij}^- = \left[\sum_{\alpha=1}^{N} w_\alpha \langle x_i \rangle_\alpha \langle x_j \rangle_\alpha \right] - \hat{x}_i^- \hat{x}_j^- \tag{3.56}$$

The updated Kalman gain K_k at the time step t_k is given by

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T [\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}]^{-1} \tag{3.57}$$

where the linearized measurement model is given by $\mathbf{H}_k = \frac{\partial \mathbf{h}}{\partial \mathbf{x}}|_{\mathbf{x}=\hat{\mathbf{x}}_k^-}$. The estimation of the state and the estimated error covariance are updated as

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k) \tag{3.58}$$

and

$$\mathbf{P}_k^+ = [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k] \mathbf{P}_k^- \tag{3.59}$$

After the measurement update, the abscissas are re-sampled to match the updated mean and covariance.

Filter smugness

The filter smugness means the error covariance matrix $\mathbf{P}_k$ becomes very small, which results in the very small gain. Under this situation the filter becomes believing that the states are well known, and refuses to incorporate the new information from the measurements [72]. The problem is particularly drastic if the noise inputs to the system and measurement noise are small [38]. Examination of the estimation algorithm shows that, as $\mathbf{P}_k$ becomes very small, the filter becomes less sensitive to the measurements from the sensor, and the estimate will fail to correct errors or diverge. Several solutions have been devised to deal with this problem [11, 38, 72], and they can be categorized into two different approaches, adaptive and non-adaptive. In this work a non-adaptive approach is employed to keeping the gain becoming too small. Instead of using the covariance $\mathbf{P}_k^-$ from DQMOM, a fixed $\bar{\mathbf{P}}$ is used. Thus, the new

equations for Kalman gain is

$$\mathbf{K}_k = \bar{\mathbf{P}}\mathbf{H}_k^T[\mathbf{H}_k\bar{\mathbf{P}}\mathbf{H}_k^T + \mathbf{R}]^{-1} \tag{3.60}$$

Appropriate fixed covariance $\bar{\mathbf{P}}$ can be found/designed through extensive simulation [38].

Update through the Unscented Kalman Filter

As demonstrated in the previous section, in the EKF update, a linearized model is required for the measurement. To eliminate the need for linearization of the measurement model, here the unscented transformation is used. The unscented Kalman filter uses $2N_s + 1$ scalar weights W_j^m and W_j^c for the mean and covariance respectively, which are calculated as

$$\begin{aligned} W_0^m &= \frac{\lambda}{n+\lambda} \\ W_0^c &= \frac{\lambda}{n+\lambda} + \left(1 - \alpha^2 + \beta\right) \\ W_i^m &= W_i^c = \frac{1}{2\left(N_s+\lambda\right)} \qquad i = 1, 2\cdots 2N_s \end{aligned} \tag{3.61}$$

where $\lambda = \alpha^2(N_s+\kappa) - N_s$ is a scaling parameter. The significance of the parameters α, β and κ can be found in [73]. The update stage of the DQMOM-UKF implementation consists of the following steps. First the prediction of the states $\hat{\mathbf{x}}_k^-$ and the covariance $\mathbf{P}_k^-$ are coming from the DQMOM. The sigma point vector at this time step is constructed by

$$\boldsymbol{\mathcal{X}} = \left[\begin{array}{ccc} \hat{\mathbf{x}}_k^- & \hat{\mathbf{x}}_k^- + \left(\sqrt{\left(N_s+\lambda\right)\mathbf{P}_k^-}\right)_i & \hat{\mathbf{x}}_k^- - \left(\sqrt{\left(N_s+\lambda\right)\mathbf{P}_k^-}\right)_{i+N_s} \end{array}\right] \tag{3.62}$$

where $i = 1, \cdots, N_s$ The predicted measurement is $\hat{\mathbf{y}}_k = \sum_{i=0}^{2N_s} W_i^m \mathcal{Y}_i$ where $\mathcal{Y}_i = \mathbf{h}\left(\mathcal{X}_i\right)$ $i = 1, 2\cdots 2N_s + 1$ where $\mathbf{h}$ is nonlinear measurement function. The innovation

covariance $\mathbf{P}_{xy}$ and the cross covariance $\mathbf{P}_{xy}$ are

$$\begin{aligned}\mathbf{P}_{yy} &= \sum_{i=0}^{2N_s} W_i^c \left(\mathcal{Y}_i - \hat{\mathbf{y}}_k\right)\left(\mathcal{Y}_i - \hat{\mathbf{y}}_k\right)^T \\ \mathbf{P}_{xy} &= \sum_{i=0}^{2N_s} W_i^c \left(\mathcal{X}_i - \hat{\mathbf{x}}_k^-\right)\left(\mathcal{Y}_i - \hat{\mathbf{y}}_k\right)^T\end{aligned} \tag{3.63}$$

The Kalman gain is $\mathcal{K}_k = \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}$. The estimation of the state and the estimated error covariance are given as

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathcal{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k) \tag{3.64}$$

$$\mathbf{P}_k^+ = \mathbf{P}_k^- - \mathcal{K}_k\mathbf{P}_{yy}\mathcal{K}_k^T \tag{3.65}$$

The measurement dynamics is linearized with a first-order Taylor series expansion about the current state estimate for EKF while unscented transformation (UT) is employed for UKF. The UT in UKF is as accurate as a second-order Taylor series expansion. Comparison of measurement process between EKF and UKF are well illustrated in reference [68].

3.4 Summary of the Chapter

The focus of this research is to investigate innovative nonlinear estimation algorithms by improving the computational efficiency associated with solving the FPE substantially. This purpose has been achieved by two different ways. First one is through the use of a finite different method. Regarding the finite difference method, an efficient and simple adaptive moving domain that can be applied to both the explicit method and the ADI method is developed to increase the computational efficiency in solving the FDE. However, the combination of ADI method and the moving domain is not enough for the high dimensional problem such as the orbit determination problem.

So, the more efficient method, DQMOM, is employed as the second method to solve the FPE. The DQMOM converts the FPE into algebraic differential equations, which makes finding the solution very efficient. Based on the DQMOM, a hybrid filtering algorithm is formed by employing the update equations from the EKF and the UKF to mitigate the "degeneracy" phenomenon observed when the update equation based on Bayes' formula is used.

CHAPTER 4

Applications of the Nonlinear Filtering Algorithms

The proposed nonlinear filtering algorithms are applied to three different applications, i) the spacecraft relative position estimation, ii) the orbit determination problem, and iii) the bearing-only tracking problem. The numerical simulation results are presented. The filtering algorithm based on the ADI method with the moving domain is applied to the relative orbit determination problem, and the DQMOM based filtering method is applied to the bearing-only tracking and the orbit determination problem. The performance of the filters is compared with existing nonlinear filtering method such as the extended Kalman filter.

All of the simulations in this chapter were done using in a workstation with the Intel® 2.33 GHz Xeon processor and coded with MatLab®.

4.1 Relative Orbit Determination

The nonlinear filtering via the solution of the FPE is applied to the relative position problem. The FPE is evaluated using the ADI method with the moving domain.

4.1.1 Relative Orbit Dynamics

The relative motion of spacecraft in a formation flying mission can be described by the Hill's equation given by Eq. (4.1) if the leader is in a near circular orbit, and the distance between the leader and the follower satellite is small [72]. The mean motion of the leader orbit is $n = \sqrt{\mu/r_l^3}$, where μ is the Earth gravitational parameter and r_l is the orbit radius of the leader satellite. The in-plane, x-y direction, and the

out-of-plane, z direction, motions are decoupled as

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 3n^2 & 0 & 0 & 0 & 2n & 0 \\ 0 & 0 & 0 & -2n & 0 & 0 \\ 0 & 0 & -n^2 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ f_x \\ f_y \\ f_z \end{bmatrix} \triangleq \mathbf{F} + \mathbf{f} \tag{4.1}$$

Assuming that the process noise (may come from the differential drag and differential J_2) is Brownian with a diffusion matrix $\mathbf{G}$ and an autocorrelation matrix $\mathbf{Q}$ as following

$$\mathbf{G} = \begin{bmatrix} \mathbf{O}_{3\times3} & \mathbf{O}_{3\times3} \\ \mathbf{O}_{3\times3} & \mathbf{I}_{3\times3} \end{bmatrix} \quad \text{and} \mathbf{Q} = \begin{bmatrix} \mathbf{O}_{3\times3} & \mathbf{O}_{3\times3} \\ \mathbf{O}_{3\times3} & \boldsymbol{\sigma}_{3\times3} \end{bmatrix} \tag{4.2}$$

where the matrix $\boldsymbol{\sigma}$ is the diagonal matrix whose diagonal members are $\left[\sigma_{\dot{x}}^2, \sigma_{\dot{y}}^2, \sigma_{\dot{z}}^2\right]$. Then, according to Eq. (3.6), the drift term can be derived as

$$-\sum_{i=1}^{6} \frac{\partial \left[p\mathbf{F}_i\right]}{\partial x_i} = -\left(\frac{\partial p}{\partial x}\dot{x} + \frac{\partial p}{\partial y}\dot{y} + \frac{\partial p}{\partial z}\dot{z} + \left(2n\dot{y} + 3n^2 x\right)\frac{\partial p}{\partial \dot{x}} - 2n\dot{x}\frac{\partial p}{\partial \dot{y}} - n^2 z\frac{\partial p}{\partial \dot{z}}\right) \tag{4.3}$$

whereas the diffusion term is

$$\frac{1}{2}\sum_{i=1}^{6}\sum_{j=1}^{6} \frac{\partial^2 \left[p\left(\mathbf{G}\mathbf{Q}\mathbf{G}^T\right)_{ij}\right]}{\partial x_i \partial x_j} = \frac{\sigma_{\dot{x}}^2}{2}\frac{\partial^2 p}{\partial \dot{x}^2} + \frac{\sigma_{\dot{y}}^2}{2}\frac{\partial^2 p}{\partial \dot{y}^2} + \frac{\sigma_{\dot{z}}^2}{2}\frac{\partial^2 p}{\partial \dot{z}^2} \tag{4.4}$$

Hence, the resulting FPE for Hill's equations becomes

$$\frac{\partial p}{\partial t} = -\left(\frac{\partial p}{\partial x}\dot{x} + \frac{\partial p}{\partial y}\dot{y} + \frac{\partial p}{\partial z}\dot{z}\right) - \left(2n\dot{y} + 3n^2 x\right)\frac{\partial p}{\partial \dot{x}} + 2n\dot{x}\frac{\partial p}{\partial \dot{y}} + n^2 z\frac{\partial p}{\partial \dot{z}} + \frac{\sigma_{\dot{x}}^2}{2}\frac{\partial^2 p}{\partial \dot{x}^2} + \frac{\sigma_{\dot{y}}^2}{2}\frac{\partial^2 p}{\partial \dot{y}^2} + \frac{\sigma_{\dot{z}}^2}{2}\frac{\partial^2 p}{\partial \dot{z}^2} \quad (4.5)$$

4.1.2 ADI method Setup

The operators for the ADI method are

$$A_1 = a_1\frac{\partial}{\partial x}, \quad A_2 = a_2\frac{\partial}{\partial y}, \quad A_3 = a_3\frac{\partial}{\partial z} \quad (4.6)$$

and

$$A_4 = a_4\frac{\partial}{\partial \dot{x}} + \frac{Q_\alpha}{2}\frac{\partial^2}{\partial \dot{x}^2}, \quad A_4 = a_5\frac{\partial}{\partial \dot{y}} + \frac{Q_\alpha}{2}\frac{\partial^2}{\partial \dot{y}^2}, \quad A_5 = a_6\frac{\partial}{\partial \dot{z}} + \frac{Q_\alpha}{2}\frac{\partial^2}{\partial \dot{z}^2} \quad (4.7)$$

where $a_1 = -\dot{x}$, $a_2 = -\dot{y}$, $a_3 = -\dot{z}$, $a_4 = -\left(2n\dot{y} + 3n^2 x\right)$, $a_5 = 2n\dot{x}$, and $a_6 = n^2 z$, respectively. So, the ADI scheme for a general 6-order system is

$$p^{t+\frac{k}{6}} = \left(\mathbf{I} - \Delta t A_k\right)^{-1} p^{t+\frac{k-1}{6}} \qquad k = 1 \ldots 6 \quad (4.8)$$

As shown in Chapter 2, after discretizing in space $(\mathbf{I} - \Delta t A_k)$ becomes a tridiagonal matrix and can be inverted very quickly using Thomas' algorithm [70]. The operators in the form of Eq. (4.8) with boundary conditions of p_{i_1} and $p_{i_{max}}$ can be

rewritten as when $a_i \geq 0$

$$\begin{bmatrix} p_{i_1}^{t+\frac{k}{N}} \\ p_{i_2}^{t+\frac{k}{N}} \\ \vdots \\ \vdots \\ p_{i_{\max-2}}^{t+\frac{k}{N}} \\ p_{i_{\max-1}}^{t+\frac{k}{N}} \end{bmatrix} = \begin{bmatrix} M & N & 0 & \cdots & \cdots & 0 \\ 0 & M & N & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & N & 0 \\ 0 & \cdots & 0 & 0 & M & N \\ 0 & \cdots & 0 & 0 & 0 & M \end{bmatrix}^{-1} \begin{bmatrix} p_{i_1}^{t+\frac{k-1}{N}} \\ p_{i_2}^{t+\frac{k-1}{N}} \\ \vdots \\ \vdots \\ p_{i_{\max-2}}^{t+\frac{k-1}{N}} \\ p_{i_{\max-1}}^{t+\frac{k-1}{N}} + a\frac{\Delta t}{\Delta \alpha} p_{i_{\max}}^{t+\frac{k-1}{N}} \end{bmatrix} \tag{4.9}$$

where $M = 1 + a_i \dfrac{\Delta t}{\Delta \alpha_i}$ and $N = -a_i \dfrac{\Delta t}{\Delta \alpha_i}$.

When $a_i < 0$

$$\begin{bmatrix} p_{i_1}^{t+\frac{k}{N}} \\ p_{i_2}^{t+\frac{k}{N}} \\ \vdots \\ \vdots \\ p_{i_{\max-2}}^{t+\frac{k}{N}} \\ p_{i_{\max-1}}^{t+\frac{k}{N}} \end{bmatrix} = \begin{bmatrix} M & 0 & 0 & \cdots & \cdots & 0 \\ L & M & 0 & 0 & \cdots & 0 \\ 0 & L & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & 0 \\ 0 & \cdots & 0 & L & M & 0 \\ 0 & \cdots & 0 & 0 & L & M \end{bmatrix}^{-1} \begin{bmatrix} p_{i_1}^{t+\frac{k-1}{N}} - a\frac{\Delta t}{\Delta \alpha} p_0^{t+\frac{k-1}{N}} \\ p_{i_2}^{t+\frac{k-1}{N}} \\ \vdots \\ \vdots \\ p_{i_{\max-2}}^{t+\frac{k-1}{N}} \\ p_{i_{\max-1}}^{t+\frac{k-1}{N}} \end{bmatrix} \tag{4.10}$$

where $L = a_i \dfrac{\Delta t}{\Delta \alpha_i}$ and $M = 1 - a_i \dfrac{\Delta t}{\Delta \alpha_i}$. When $a_i \geq 0$ and $a_i < 0$ the operators in the form of Eq. (4.8) with boundary conditions of p_{j_1} and $p_{j_{max}}$ can be rewritten as

Eq. (4.11) and Eq. (4.12) accordingly

$$\begin{bmatrix} p_{j_1}^{t+\frac{k+1}{N}} \\ p_{j_2}^{t+\frac{k+1}{N}} \\ \vdots \\ \vdots \\ p_{j_{\max-2}}^{t+\frac{k+1}{N}} \\ p_{j_{\max-1}}^{t+\frac{k+1}{N}} \end{bmatrix} = \begin{bmatrix} M & N & 0 & \cdots & \cdots & 0 \\ -L & M & N & 0 & \cdots & 0 \\ 0 & -L & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & N & 0 \\ 0 & \cdots & 0 & -L & M & N \\ 0 & \cdots & 0 & 0 & -L & M \end{bmatrix}^{-1} \begin{bmatrix} p_{j_1}^{t+\frac{k}{N}} + a \cdot p_{j0}^{t+\frac{k}{N}} \\ p_{j_2}^{t+\frac{d-1}{N}} \\ \vdots \\ \vdots \\ p_{j_{\max-2}}^{t+\frac{k}{N}} \\ p_{j_{\max-1}}^{t+\frac{k}{N}} + b \cdot p_{j\max}^{t+\frac{k}{N}} \end{bmatrix} \tag{4.11}$$

where $L = \dfrac{Q_j \Delta t}{2\Delta\dot{\alpha}_j^2}$, $M = 1 + b_j \dfrac{\Delta t}{\Delta\dot{\alpha}_j} + \dfrac{Q_j \Delta t}{\Delta\dot{\alpha}_j^2}$, and $N = -\left(b_j \dfrac{\Delta t}{\Delta\dot{\alpha}_j} + \dfrac{Q_j \Delta t}{2\Delta\dot{\alpha}_j^2} \right)$

$$\begin{bmatrix} p_{j1}^{t+\frac{k+1}{N}} \\ p_{j2}^{t+\frac{k+1}{N}} \\ \vdots \\ \vdots \\ p_{j\max-2}^{t+\frac{k+1}{N}} \\ p_{j\max-1}^{t+\frac{k+1}{N}} \end{bmatrix} = \begin{bmatrix} M & -N & 0 & \cdots & \cdots & 0 \\ L & M & -N & 0 & \cdots & 0 \\ 0 & L & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & -N & 0 \\ 0 & \cdots & 0 & L & M & -N \\ 0 & \cdots & 0 & 0 & L & M \end{bmatrix}^{-1} \begin{bmatrix} p_{j1}^{t+\frac{k}{N}} - a \cdot p_{j0}^{t+\frac{k}{N}} \\ p_{j2}^{t+\frac{k}{N}} \\ \vdots \\ \vdots \\ p_{j\max-2}^{t+\frac{k}{N}} \\ p_{j\max-1}^{t+\frac{k}{N}} + c \cdot p_{j\max}^{t+\frac{k}{N}} \end{bmatrix} \tag{4.12}$$

where $L = b_j \dfrac{\Delta t}{\Delta\dot{\alpha}_j} - \dfrac{Q_j \Delta t}{\Delta\dot{\alpha}_j^2}$, $M = 1 - b_j \dfrac{\Delta t}{\Delta\dot{\alpha}_j} + \dfrac{Q_j \Delta t}{\Delta\dot{\alpha}_j^2}$ and $N = \dfrac{Q_j \Delta t}{2\Delta\dot{\alpha}_j^2}$. As mentioned before, all the matrix inversions in Eq. (4.9) through Eq. (4.12) involve only tri-diagonal matrices.

4.1.3 Measurement PDF

The measurement for the relative orbit determination problem is assumed to be the range between spacecrafts (Eq. 4.13), the azimuth angle (Eq. 4.14), and the elevation angle (Eq. 4.15).

$$r = \sqrt{x^2 + y^2 + z^2} \tag{4.13}$$

$$\theta = \tan^{-1}\left(\frac{y}{x}\right) \tag{4.14}$$

$$\phi = \sin^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) = \sin^{-1}\left(\frac{z}{r}\right) \tag{4.15}$$

The probabilistic information via the measurement is captured by Eq. (3.26). The resulting PDF is the Gaussian PDF as the function of the measurements r, θ, and ϕ. To implement the Gaussian PDF in Bayes' formula Eq. (3.25), it is necessary to perform the transformation of the random variable to obtain the Gaussian PDF as function of x, y and z [38].

$$p\left(\mathbf{y}\middle| x, y, z\right) = p\left(\mathbf{y}\middle| r, \theta, \phi\right) \left|\mathbf{J}\left(x, y, z\right)\right| \tag{4.16}$$

where $\left|\mathbf{J}\left(x, y, z\right)\right|$ is the Jacobian determinant of the measurement functions, and the Jacobian determinant of the measurement functions is

$$\left|\mathbf{J}\left(x, y, z\right)\right| = \frac{\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2 + z^2}\left(x^2 + y^2 + z^2\right)} \tag{4.17}$$

Eq. (4.16) becomes

$$p\left(\mathbf{y}\middle| x, y, z\right) = p\left(\mathbf{y}\middle| r, \theta, \phi\right) \frac{\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2 + z^2}\left(x^2 + y^2 + z^2\right)} \tag{4.18}$$

Therefore, the measurement PDF, required for the Bayes' formula (Eq. (3.25)) can be acquired by evaluating the Eq. (4.18) for all values of $\dot{x}$, $\dot{y}$ and $\dot{z}$

4.1.4 Partitioning Measurement PDF

Even with the help from the proposed adaptive moving domain technique and the ADI method, solving the FPE numerically is still too expensive due to the curse of dimensionality [18]. Luckily, the Hill's equation has decoupled in-plane and the out-of-plane motions. Therefore, the conditional state PDF in Bayes' formula, $p(\mathbf{x}|\mathbf{Y}) = p(x, y, z, \dot{x}, \dot{y}, \dot{z}|\mathbf{Y})$ can be partitioned into disjoint PDFs as $p(x, y, \dot{x}, \dot{y}|\mathbf{Y})$ and $p(z, \dot{z}|\mathbf{Y})$. Consequently, these PDFs can be evaluated separately. Thus, the computational cost can be reduced even further. However, the measurements are functions of both the in-plane and the out-of-plane coordinates. So the following assumption has to be made for this problem

$$p\left(\mathbf{x_1}, \mathbf{x_2} \left| \mathbf{Y}\left(t_{k-1}\right)\right.\right) \approx p\left(\mathbf{x_1} \left| \mathbf{Y}\left(t_{k-1}\right)\right.\right) p\left(\mathbf{x}_2 \left| \mathbf{Y}\left(t_{k-1}\right)\right.\right) \tag{4.19}$$

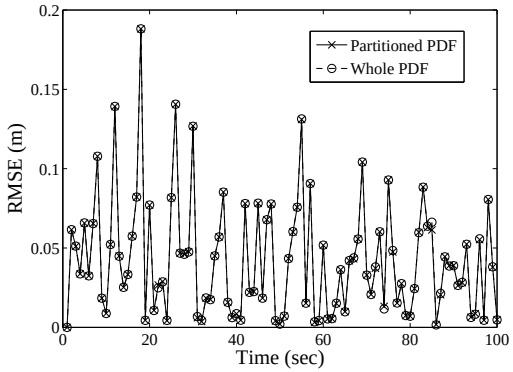
where $\mathbf{x}_1 = [x, y, \dot{x}, \dot{y}]^T$ is the states of the in-plane motion, $\mathbf{x}_x = [z, \dot{z}]^T$ is the states of the out-of-plane motion. Therefore, the Bayes' formula becomes

$$p\left(\mathbf{x}_k \left| \mathbf{Y}_{k-1}\right.\right) = \frac{p\left(\mathbf{x_1} \left| \mathbf{Y}_{k-1}\right.\right) p\left(\mathbf{x}_2 \left| \mathbf{Y}_{k-1}\right.\right) p\left(\mathbf{y}_k \left| \mathbf{x}_k\right.\right)}{\int p\left(\xi \left| \mathbf{Y}_{k-1}\right.\right) p\left(\mathbf{y}_k \left| \xi\right.\right) d\xi} \tag{4.20}$$

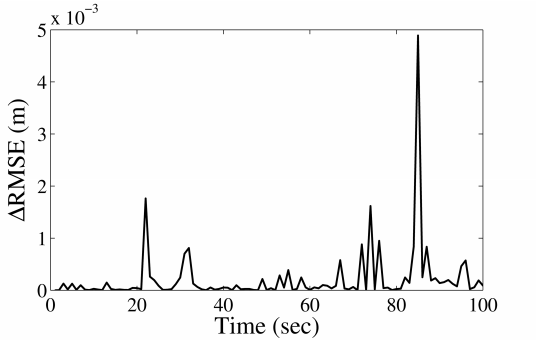
The difference between the true conditional PDF and the approximated one can be calculated as

$$\Delta\varepsilon = p\left(\mathbf{x_1}, \mathbf{x_2} \left| \mathbf{Y}_{k-1}\right.\right) - p\left(\mathbf{x_1} \left| \mathbf{Y}_{k-1}\right.\right) p\left(\mathbf{x}_2 \left| \mathbf{Y}_{k-1}\right.\right) \tag{4.21}$$

If Eq. (4.19) is a suitable assumption for this problem, $\Delta\varepsilon$ given in Eq. (4.21) should be small. Specifically, the estimation results from two estimators. One uses the whole PDF, $p(\mathbf{x}_1, \mathbf{x}_2|\mathbf{Y}_{k-1}$ and the other uses the partitioned PDFs, $p(\mathbf{x}_1|\mathbf{Y})$ and $p(\mathbf{x}_2|\mathbf{Y})$. Under the same simulation conditions, there should not be any significant difference in their estimation accuracy. Here, a comparison is made in terms of the root mean square errors (RMSEs) between these two estimators. Only the comparison for the



(a) Comparison of the RMSE in x



(b) The absolute difference between the RMSEs in x using the partitioned PDF and the whole PDF

Figure 4.1: Comparison of Estimations

coordinate has been illustrated as shown in Fig. 4.1(a)-4.1(b) which is generated using the same condition that will be used in numerical simulation and result section to be described later; however note that similar results have been achieved for the other coordinates.

It is shown in Fig. 4.1(b) that the difference is on the order of 10^{-3}, which is well below the order of the RMSE (10^{-1}-10^{-2}). This comparison was conducted for several different cases and similar results were obtained. This shows that the effect of on the quality of estimation is not significant and it is safe to assume that Eq. (4.19) is valid for the relative orbit position estimation problem. However, a similar comparison is encouraged to be performed such that the partition is valid for any specific problems.

In terms of the computational time, the case using the whole PDF in the estimation requires approximately 47 seconds while that of the partitioned PDF is 3 seconds. Therefore, the computational cost when whole PDF is using is about sixteen times higher. Therefore, balancing between accuracy (not much compromise) and speed, the partitioned PDF method will be used in this paper.

4.1.5 Simulation Setup

Two satellites are assumed in the simulated formation flying mission. The orbit parameters for the leader satellite are the semi-major axis $a = 6978\ km$, eccentricity $e = 0$, inclination $i = 97.8^o$, argument of periapsis $\omega = 97.8^o$, longitude of the ascending node $\Omega = 100.7^o$, and true anomaly $f = 100.7^o$. The follower orbit is designed to satisfy Eq. (4.22) in order to achieve a bounded motion centered around the leader satellite location [72]

$$\dot{y}_0 = -2nx_0 \text{ and } y_0 = \frac{2\dot{x}_0}{n} \tag{4.22}$$

where $x_0 = 86.52\ m$ and $\dot{x}_0 = 0.5336\ m/s$ are the initial relative position and the velocity in the local vertical coordinate, and y_0 and $\dot{y}_0$ are the initial relative position and the velocity in local horizontal coordinate. The initial relative position and the velocity in the out-of-plane coordinate are $z_0 = 92.65\ m$ and $\dot{x}_0 = 0.3003\ m/s$ The initial relative distance between two satellites is about 1 km. The process noise is assumed to be a Gaussian with a zero mean and a power spectral density of $(10^{-3}\ m/s)^2$ The measurement noise in the relative distance is assumed to be zero mean white Gaussian with variance of $(0.1\ m)^2$, and the measurement noise in both the azimuth angle and the elevation angle are assumed to be zero mean white Gaussian with variance of $(10\ arcsecond)^2$. Measurement update frequency is 1 Hz. 1000 seconds are simulated. For comparison, the same setup is applied to the EKF.

Three cases with different measurement update frequencies are simulated. When the measurement is updated at every 1 second (1 Hz), the setup for the ADI method shown in Table 4.1 is used.

The simulation setup when the measurement updates are at every 5 seconds (0.2 Hz) and 10 seconds (0.1 Hz) are presented in Table 4.2. The size of the domain for less measurement update case is increased noticeably in x, y, and z since the time between

Table 4.1: Grid and domain size used for the case of 1 Hz measurement update

	x (m)	y (m)	z (m)	$\dot{x}$ (m/s)	$\dot{y}$ (m/s)	$\dot{z}$ (m/s)
Grid size	0.12	0.1	0.07	0.02	0.02	0.02
Domain size	2.16	1.8	1.12	0.24	0.24	0.24

Table 4.2: Grid and domain size setup for the cases of 0.1 and 0.2 Hz measurement updates

	x (m)	y (m)	z (m)	$\dot{x}$ (m/s)	$\dot{y}$ (m/s)	$\dot{z}$ (m/s)
Grid size	0.3	0.2	0.2	0.02	0.02	0.02
Domain size	7.2	4.0	4.0	0.24	0.24	0.24

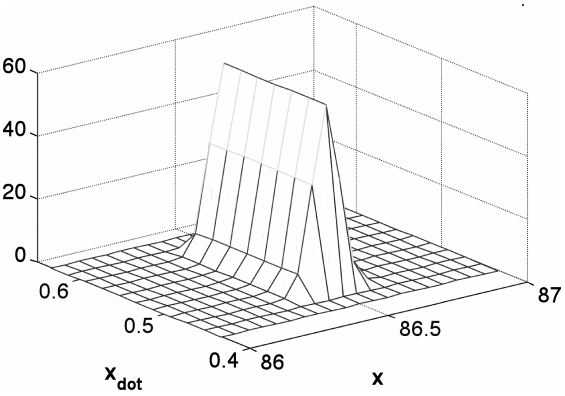
measurement updates longer, thus it requires larger space for PDF to propagate.

The initial PDF for the FPE can be obtained using Eq. (4.18) in each direction based on the statistical characteristics of the sensor. Figure 4.2 shows an initial PDF of the in-plane motion (a and b) and the out-of-plane motion (c). It can be seen that the initial appearances of the PDFs from the x, y, and z directions are Gaussian since the sensor noise is Gaussian white noise while they look like an uniform distribution from other directions due to the absence of the sensor measurement.

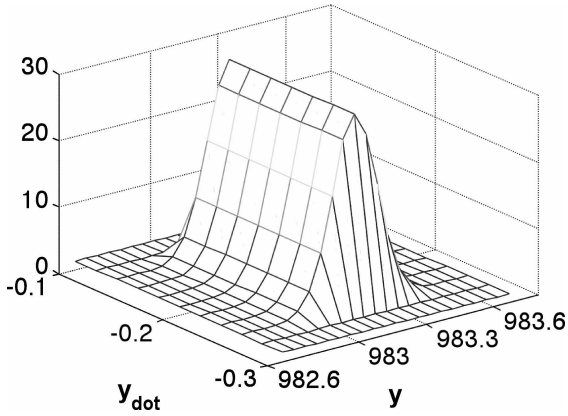
The initial PDFs are evolved using FPE and the evolved PDFs are illustrated in Figure 4.3. The shape of the PDFs in $\dot{x}$, $\dot{y}$, and $\dot{z}$ directions are transformed from the initial uniform distribution to the Gaussian-like distribution.

4.1.6 Simulation Results

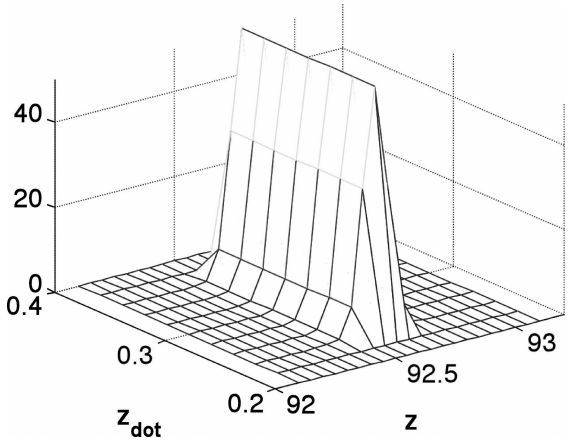
The results of the forty Monte Carlo simulations are presented in this section. The root mean square error (RMSE) of the position estimation in the x direction for different measurement update frequencies is shown in Fig. 4.4. Three methods, EKF, nonlinear filtering with the ADI method with the moving domain (denoted as 'ADI FPE'), and nonlinear filtering with the conventional explicit method with the moving domain (denoted as 'EXP FPE')) are compared here. It can be seen that the FPE



(a) PDF in $x(m)$ and $\dot{x}(m/s)$ coordinate



(b) PDF in $x(m)$ and $\dot{x}(m/s)$ coordinate



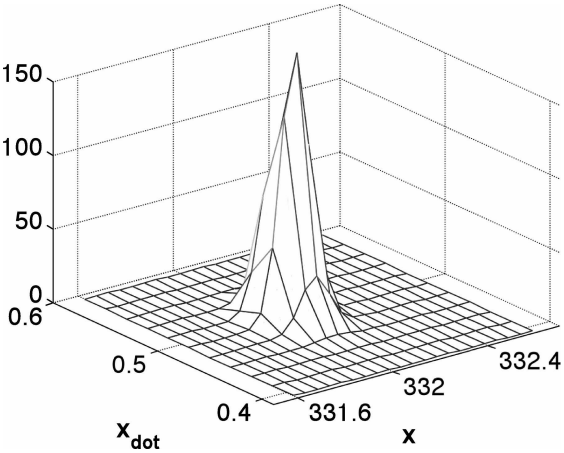
(c) PDF in $x(m)$ and $\dot{x}(m/s)$ coordinate

Figure 4.2: Initila PDFs

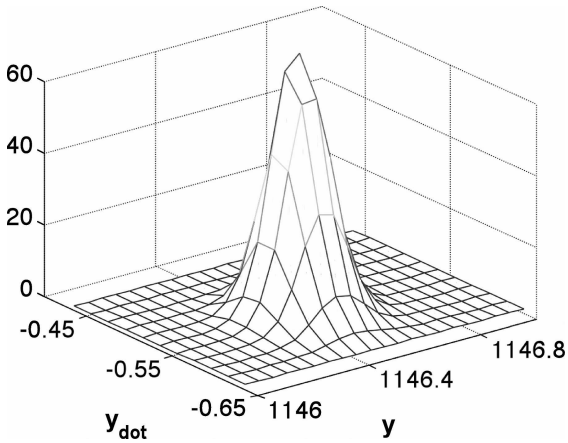
based nonlinear filer performs similarly to that of the EKF when the measurement update is frequent (e.g. 1 Hz). However, as the measurement update frequency decreases, the performance of the EKF degrades whereas those of the ADI method are consistent. The same conclusion can be made for the other two directions (y and z) and the simulation plots are neglected here.

Table 4.3 indicates that the average RMSE of the EKF increases (although still acceptable) while the average RMSE of the estimator using the ADI method is insensitive to the less frequent update.

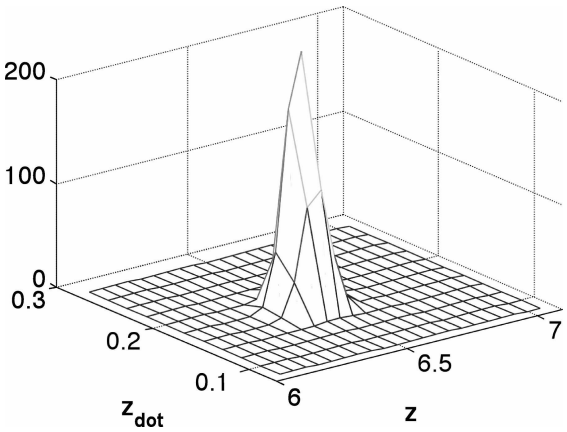
The root mean square error (RMSE) of the velocity estimation in each direction is presented in Fig. 4.5 when the measurement is updated at 0.2 Hz. It can be seen in



(a) PDF in $x(m)$ and $\dot{x}(m/s)$ coordinate



(b) PDF in $x(m)$ and $\dot{x}(m/s)$ coordinate

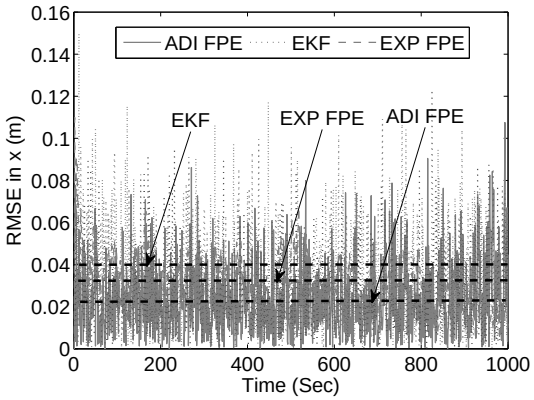


(c) PDF in $x(m)$ and $\dot{x}(m/s)$ coordinate

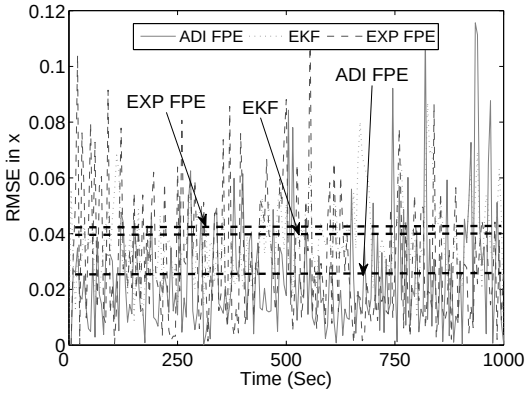
Figure 4.3: Evolved PDFs

these figures that the performance in the velocity estimation is very similar for both EKF and ADI methods. A similar argument can be made for the other two cases and the plots are neglected here.

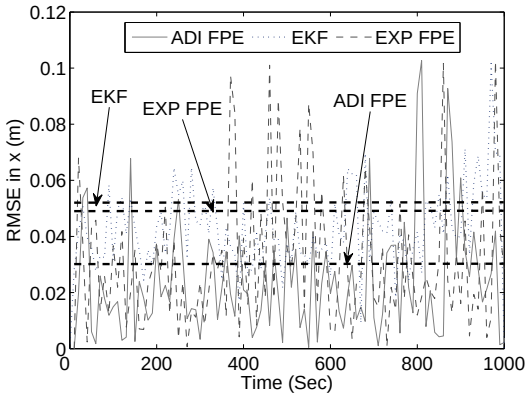
Although in general the computational cost for the ADI method is still intensive as compared with the EKF as shown in Table 4.4, it is found that this method is nearly five times faster than the conventional explicit method with a better performance as shown in Table 4.3. Furthermore, it is known that in general, the execution speed of the program coded with MATLAB® is much slower than the one in C. Therefore, the real time realization of this work could be possible with the ADI method in conjunction with the adaptive moving domain technique which decreases the size of



(a) 1 Hz



(b) 0.2 Hz



(c) 0.1 Hz

Figure 4.4: Root mean square errors in x direction

the necessary domain and the partition of PDF which lowers the problem dimension.

The accuracy of the estimation method used in this dissertation is limited by the accuracy of the numerical method. For example, the accuracies of the ADI method and the explicit method are only $O(\Delta t)$ and $O(\Delta x)$. A higher accuracy, $O(\Delta t^2)$ in time, can be achieved by using the D'Yaknove scheme [70] instead of the implicit Euler method. There also exists an up-wind scheme with an accuracy of $O(\Delta x^2)$, but the computational cost will be much higher [80].

Table 4.3: Average RMSE (in m)

Update delay (s)	Direction	EKF	EXP FPE	ADI FPE
1	x	0.04	0.035	0.025
	y	0.03	0.03	0.03
	z	0.02	0.02	0.02
5	x	0.042	0.04	0.025
	y	0.055	0.045	0.03
	z	0.03	0.022	0.02
10	x	0.043	0.04	0.025
	y	0.058	0.035	0.03
	z	0.04	0.02	0.02

4.2 Bearing-only Tracking

In this section, a standard bearing only tracking problem is used to demonstrate the effects of the modified DQMOM nonlinear filtering technique. The performance will be compared among the EKF, UKF, DQMOM-EKF, and DQMOM-UKF techniques.

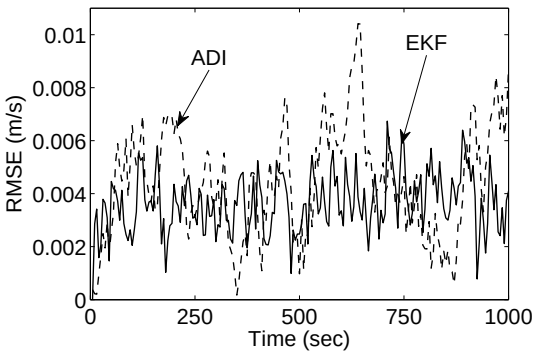
4.2.1 Dynamics and Measurement

A simplified version of the passive bearing only tracking problem is adopted from [8, 68] as the example. The motion of the sensor platform is governed by $x_p = 4t$ and $y_p = 20$, whereas the motion of the target is governed by

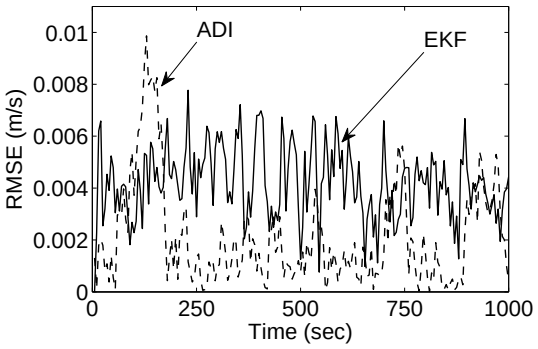
$$\frac{dx(t)}{dt} = \frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t) \tag{4.23}$$

Table 4.4: Computational cost (in seconds)

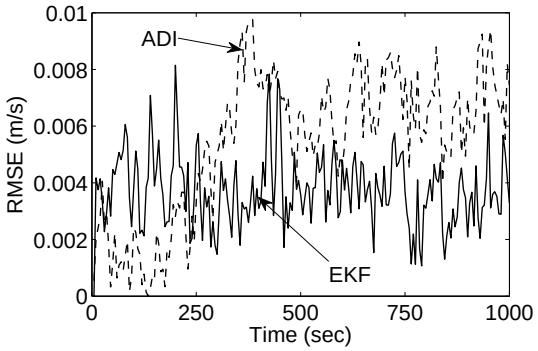
Update delays (s)	EKF	EXP FPE	ADI FPE
1	1×10^{-5}	11	3
5	2×10^{-5}	43	8.3
10	1×10^{-5}	80	14.6



(a) RMSE in $\dot{x}$



(b) RMSE in $\dot{y}$



(c) RMSE in $\dot{z}$

Figure 4.5: Root mean square errors in velocity estimation

where the initial condition of the target is $x(t_0) = [80, 1]^T$, and the process noise is a Gaussian with a zero mean and a covariance of $Q = 10^{-2}$. The measurement model is given by

$$y(t_k) = \tan^{-1} \frac{y_p(t_k)}{x_1(t_k) - x_p(t_k)} + v(t_k) \tag{4.24}$$

where the sensor noise $v(t_k)$ is assumed to have a zero mean with a variance of $R = (4^o)^2$.

4.2.2 Simulation Setup

500 Monte Carlo simulation runs are used to compare the performance achieved from the EKF, UKF, DQMOM-EKF and DQMOM-UKF filters. The random number

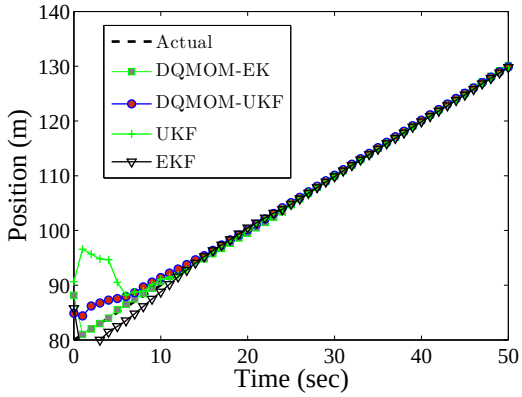
generation was reseeded with clock to ensure there is no correlation between the runs. The filters are set to make the estimation in each second regardless of the measurement update frequency. Two different measurement frequencies, 1Hz and 0.2 Hz are used to see the performance of the filters when the measurement is absent for a significant duration. The detail information regarding the UKF and the EKF setup can be found from Ref. [8], and the moment constraints for DQMOM, given in Eq. (3.49) are selected to be

$$\begin{aligned} k1 &= [0\ 0\ 1\ 0\ 1\ 2\ 3\ 2\ 1\ 0\ 1\ 4] \\ k2 &= [0\ 1\ 0\ 2\ 1\ 0\ 0\ 1\ 2\ 3\ 4\ 1] \end{aligned} \tag{4.25}$$

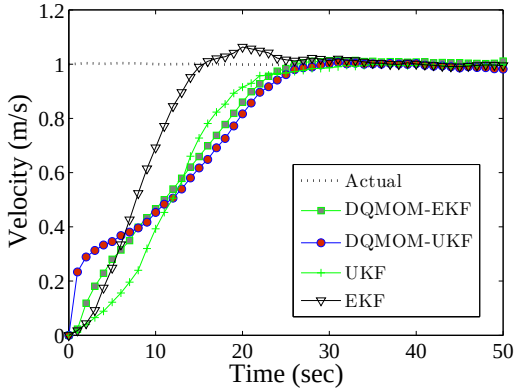
4.2.3 Simulation Results

The next two simulations are used to compare the performance achieved through different nonlinear filtering techniques. First the measurement update and the estimation frequent are set to 1 Hz. Results are presented in Fig. 4.6. The DQMOM-EKF and the DQMOM-UKF perform as good as EKF in terms of the rate of convergence and accuracy in the position estimation. However, the advantage in position estimation of the DQMOM-EKF and the DQMOM-EKF can be seen from the root-mean-square-error (RMSE) performance as demonstrated in Fig. 4.6(c). As shown in Fig. 4.6(b), the performances of all the filters are similar in the velocity estimation. Fig. 4.6(b) shows that the RMSE of both the DQMOM-UKF and the DQMOM-EKF are comparable to the filter that they borrowed the update mechanism.

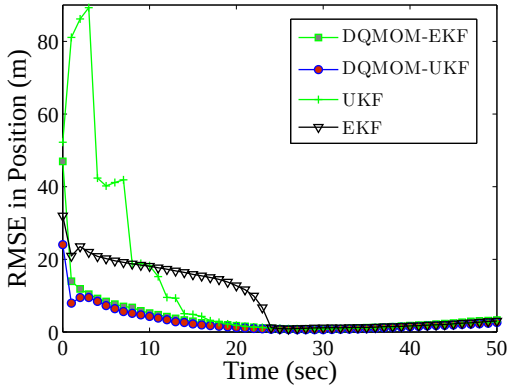
For the second simulation, the measurement update rate is reduced to 0.2 Hz while the filter is still making estimates at the rate of 1Hz. Regarding the estimation of position, the DQMOM-UKF approach shows the best performance in terms of the rate of convergence and accuracy. The result obtained using both the DQMOM-EKF approach and the EKF are similar. The RMSE in the position estimation from both



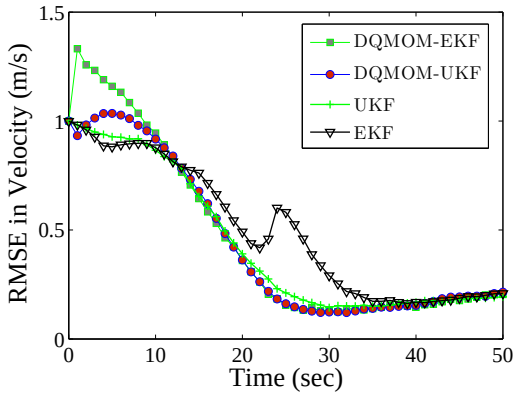
(a) Mean Position



(b) Mean Velocity

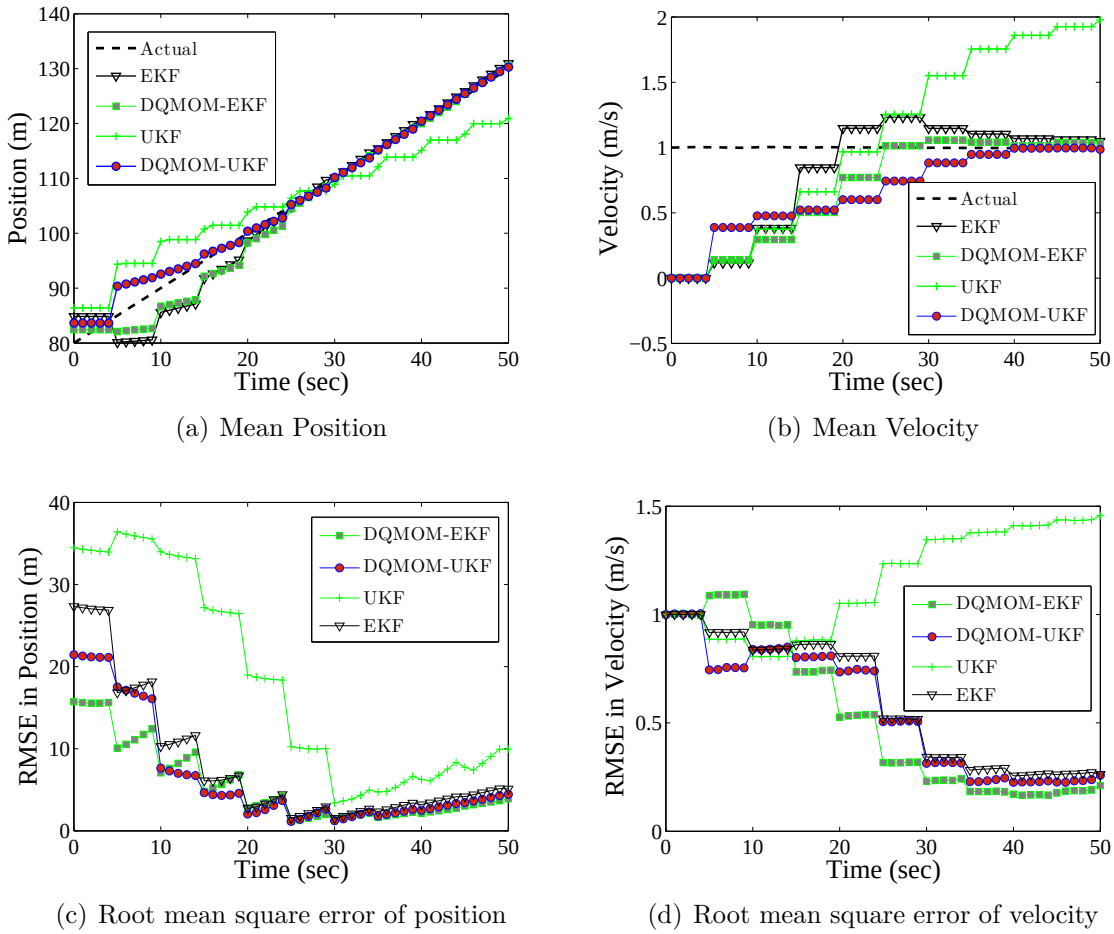


(c) Root mean square error of position



(d) Root mean square error of velocity

Figure 4.6: Estimation comparison with a measurement update frequency of 1Hz



(a) Mean Position

(b) Mean Velocity

(c) Root mean square error of position

(d) Root mean square error of velocity

Figure 4.7: Estimation comparison with a measurement update frequency of 0.2Hz

the DQMOM-EKF and the DQMOM-UKF approach are better than those obtained from the EKF and UKF (Fig.4.7(c)). For the velocity estimation, the DQMOM-EKF is showing the best result while in the performance of the RMSE these two DQMOM filters works better than the UKF and EKF in terms of convergence speed.

The computational time for running the 50-second simulation for each of the filters are list in Table 4.5. In addition to the case we mentioned before, the case with 0.5 Hz update rate has been considered. The proposed filtering algorithms are slower than EKF and UKF but are much faster than the finite difference method. It is worthwhile to notice that the same kind of simulations could take up to a minute with finite difference type methods, such as the Alternative Directional Implicit method

Table 4.5: Computational cost (in seconds)

Update delays (s)	EKF	UKF	DQMOM-EKF	DQMOM-UKF
1	0.015	0.58	4.15	4.20
5	0.015	0.60	4.22	4.20
10	0.014	0.58	4.13	4.25

with adoptive moving domain scheme. Also base on the data shown in the table, it can be found that the computational cost for 50 seconds estimation is only 4.2 seconds, which is fast enough for real-time estimation.

4.3 Orbit Determination

4.3.1 Keplerian equation of motion

The equations of motion that governs the motion of a satellite in a low earth orbit is [72]:

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r} + \boldsymbol{a}_G + \boldsymbol{a}_D \tag{4.26}$$

where μ and $\boldsymbol{r} = [x, y, z]^T$ are the gravitational parameter and the position vector, respectively. The scalar r is the magnitude of $\boldsymbol{r}$, i.e., $r = \sqrt{x^2 + y^2 + z^2}$, and $\boldsymbol{a}_D$ represents the drag due the earth atmosphere and is proportional to the atmospheric density ρ and the square of the velocity relative to the atmosphere.

$$\boldsymbol{a}_D = -\frac{1}{2B_c}\rho v_{rel}^2 \quad \text{and} \quad B_c = \frac{m_s}{C_d A}. \tag{4.27}$$

where B_c is ballistic coefficient, and m_s, C_d, and A are the mass of a satellite, the drag coefficient, and the projected area of a satellite normal to its flight path, respectively.

The earth is not a spherically symmetric body but bulged at the equator, and is also generally asymmetric. As a result, the gravitational field around the earth is not isometric. The $\boldsymbol{a}_G$ is the perturbation due to this uneven gravitational field. In this

work, $\boldsymbol{a}_D$ and $\boldsymbol{a}_G$ are considered as a part of process noise to the system. So, the nominal process equation used in this work is

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r} + \boldsymbol{w}_t \tag{4.28}$$

where $\boldsymbol{w}_t$ is a white Gaussian noise process with $E[\boldsymbol{w}_t \boldsymbol{w}_\tau^T] = \boldsymbol{Q}_t \delta(t-\tau)$

4.3.2 Fokker-Planck equation of the Keplerian equation

The state-space form of the Keplerian equation of motion with process noise is

$$\begin{aligned}
\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -\frac{\mu}{r^3} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{\mu}{r^3} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{\mu}{r^3} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{w}_t \\
&\triangleq \boldsymbol{F}\boldsymbol{x} + \boldsymbol{G}\boldsymbol{w}_t
\end{aligned} \tag{4.29}$$

where $\boldsymbol{w}_t$ is white Gaussian noise process with $E[\boldsymbol{w}_t \boldsymbol{w}_\tau^T] = \boldsymbol{Q}_t \delta(t-\tau)$

The corresponding Fokker-Planck equation can be found by substituting Eq. (4.29) into Eq. (3.6) as

$$\begin{aligned}
\frac{\partial p}{\partial t} = &-\left(\frac{\partial p}{\partial x}\dot{x} + \frac{\partial p}{\partial y}\dot{y} + \frac{\partial p}{\partial z}\dot{z}\right) + \frac{\mu}{r^3}\left(\frac{\partial p}{\partial \dot{x}}x + \frac{\partial p}{\partial \dot{y}}y + \frac{\partial p}{\partial \dot{z}}z\right) \\
&+ \frac{Q_4}{2}\frac{\partial^2 p}{\partial \dot{x}^2} + \frac{Q_5}{2}\frac{\partial^2 p}{\partial \dot{y}^2} + \frac{Q_6}{2}\frac{\partial^2 p}{\partial \dot{z}^2}
\end{aligned} \tag{4.30}$$

where Q_i, $i = 4, 5, 6$ are the last three diagonal members of $[\boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^T]$ from Eq. (3.6)

4.3.3 Measurement Model

As shown in Fig. 4.8, the inertial position vector of a satellite, $\boldsymbol{r} = [x\ y\ z]^T$ can be written as the sum of the range vector and the radar site position vector [72] as

$$\boldsymbol{\rho} = \boldsymbol{r} - \boldsymbol{R}_s \tag{4.31}$$

where $\boldsymbol{R}_s$ is the position of the sensor, and $\boldsymbol{\rho} = [\rho_e, \rho_e, \rho_n]^T$ is the position vector of a satellite relative to the local/sensor coordinate (the Topocentric-Horizon coordinate) as

$$\boldsymbol{\rho} = \rho_u \hat{\boldsymbol{u}} + \rho_e \hat{\boldsymbol{e}} + \rho_n \hat{\boldsymbol{n}} \tag{4.32}$$

wherein the subscriptions, u, e, and n stand for "zenith", "east", and "north", respectively. For the sensor position vector $\boldsymbol{R}_s$, it is advisable to account for the precise

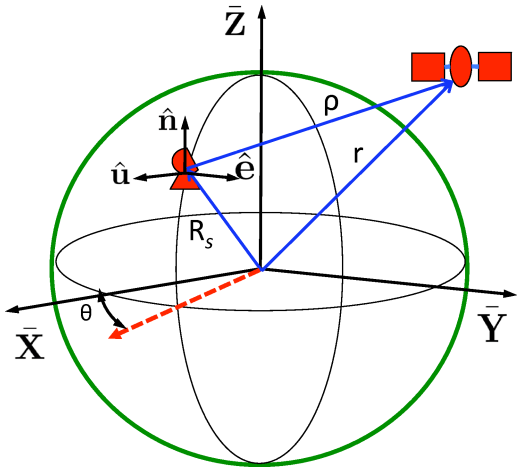


Figure 4.8: Geometry of Earth observation of a satellite

shape of the Earth to avoid large errors [72]. $\boldsymbol{R}_s$ in the geocentric inertia coordinate accounted for the Earth's equatorial bulge and its magnitude can be found from Ref. [61] as

$$\boldsymbol{R}_s = r_\delta \cos\theta \mathbf{I} + r_\delta \sin\theta \mathbf{J} + r_k \mathbf{K} \qquad \text{and} \qquad ||\boldsymbol{R}_s|| = \sqrt{r_\delta^2 + r_k^2} \tag{4.33}$$

where r is the magnitude of the position vector $\boldsymbol{r}$, θ is the sidereal time at the sensor location (local sidereal time), and r_δ and r_k can be calculated using

$$r_\delta = \left[\frac{R_\oplus}{\sqrt{1-e_\oplus^2 \sin\lambda}} + H\right]\cos\lambda \qquad \text{and} \qquad r_k = \left[\frac{R_\oplus(1-e_\oplus^2)}{\sqrt{1-e_\oplus^2 \sin\lambda}} + H\right]\sin\lambda \quad (4.34)$$

where λ is the geodesic latitude of the sensor location, and $R_\oplus = 6378.1363km$ and $e_\oplus = 0.081819221456$ are the mean equatorial radius of the Earth and the eccentricity of the Earth, respectively. H is the local elevation above the sea level.

The local sidereal time θ is the angle between the x-axis of the geocentric-equatorial coordinate (vernal equinox direction) and the longitude of the sensor location on the Earth as shown in Fig. 4.8, which can be expressed as the sum of the Greenwich sidereal time and the geographical longitude of the location.

$$\theta = \theta_{GMT} + \phi \quad (4.35)$$

where θ_{GMT} is the Greenwich sidereal time which is the angle between the Greenwich Meridian and the x-axis of the geocentric-equatorial coordinate, and ϕ is the geographical longitude of the sensor location. There are two different ways to calculate the local sidereal time (LST). The details regarding these methods are can be found in Appendix A.

The position vector $\boldsymbol{\rho}$ in the inertial coordinate frame is given by

$$\boldsymbol{\rho}^{(I)} = \begin{bmatrix} x - ||\boldsymbol{R}_s|| \cos\lambda\cos\theta \\ y - ||\boldsymbol{R}_s|| \cos\lambda\sin\theta \\ z - ||\boldsymbol{R}_s|| \sin\lambda \end{bmatrix} \quad (4.36)$$

The conversion from the inertial to the Topocentric-Horizon coordinate is given by

the following rotation matrix

$$\boldsymbol{C} = \begin{bmatrix} \cos\lambda & 0 & \sin\lambda \\ 0 & 1 & 0 \\ -\sin\lambda & 0 & \cos\lambda \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.37}$$

So the sensor position vector in the Topocentric-Horizon coordinate can be presented in terms of the components of the position vector $\boldsymbol{r}$ as

$$\boldsymbol{\rho} = \begin{bmatrix} \rho_u \\ \rho_e \\ \rho_n \end{bmatrix} = \begin{bmatrix} \cos\lambda\cos\theta & \cos\lambda\sin\theta & \sin\lambda \\ -\sin\theta & \cos\theta & 0 \\ -\sin\lambda\cos\theta & -\sin\lambda\sin\theta & \cos\lambda \end{bmatrix} \begin{bmatrix} x - ||\boldsymbol{R}_s||\cos\lambda\cos\theta \\ y - ||\boldsymbol{R}_s||\cos\lambda\sin\theta \\ z - ||\boldsymbol{R}_s||\sin\lambda \end{bmatrix} \tag{4.38}$$

The measurements are range, azimuth and elevation. The range ρ can be found from

$$\|\boldsymbol{\rho}\| = \sqrt{\rho_u^2 + \rho_e^2 + \rho_n^2} \tag{4.39}$$

The azimuth and elevation angles are

$$az = \tan^{-1}\left(\frac{\rho_e}{\rho_n}\right) \qquad el = \tan^{-1}\left(\frac{\rho_u}{\sqrt{\rho_e^2 + \rho_n^2}}\right) \tag{4.40}$$

The measurement model for the filter can be written in vector form as

$$\mathbf{H} = \left[\sqrt{\rho_u^2 + \rho_e^2 + \rho_n^2} \;\; \tan^{-1}\left(\frac{\rho_e}{\rho_n}\right) \;\; \tan^{-1}\left(\frac{\rho_u}{\sqrt{\rho_e^2 + \rho_n^2}}\right)\right]^T \tag{4.41}$$

To be applied to the EKF measurement update equations, this needs to be linearized, and the linearization can be done by calculating Jacobian of the measurement model $\mathbf{H}$ by

$$\frac{\partial \mathbf{H}}{\partial \mathbf{x}} = \left[\frac{\partial \mathbf{H}}{\partial \boldsymbol{r}} \;\; \frac{\partial \mathbf{H}}{\partial \dot{\boldsymbol{r}}}\right] = \left[\frac{\partial \mathbf{H}}{\partial \boldsymbol{r}} \;\; \mathbf{0}_{3\times 3}\right] \tag{4.42}$$

where $\mathbf{x} = [x\ y\ z\ \dot{x}\ \dot{y}\ \dot{z}]$ is the state vector from Eq. (4.29), and $\dfrac{\partial \mathbf{H}}{\partial \boldsymbol{r}}$ can be calculated using the chain rule

$$\frac{\partial \mathbf{H}}{\partial \boldsymbol{r}} = \frac{\partial \mathbf{H}}{\partial \boldsymbol{\rho}} \frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\rho}^{(I)}} \frac{\partial \boldsymbol{\rho}^{(I)}}{\partial \boldsymbol{r}} \tag{4.43}$$

Since $\boldsymbol{r} = [x\ y\ z]^T$ and Eq. (4.36), the last term of the partial derivative is $\dfrac{\partial \boldsymbol{\rho}^{(I)}}{\partial \boldsymbol{r}} = \mathbf{I}$. The second term can be written as

$$\frac{\partial \boldsymbol{\rho}}{\partial \boldsymbol{\rho}^{(I)}} = \frac{\partial \boldsymbol{\rho}^{(I)} \boldsymbol{C}}{\partial \boldsymbol{\rho}^{(I)}} = \frac{\partial \boldsymbol{C}}{\partial \boldsymbol{\rho}^{(I)}} \boldsymbol{\rho}^{(I)} + \frac{\partial \boldsymbol{\rho}^{(I)}}{\partial \boldsymbol{\rho}^{(I)}} \boldsymbol{C} = \boldsymbol{C} \tag{4.44}$$

So the linearized measurement model is $\dfrac{\partial \mathbf{H}}{\partial \boldsymbol{r}} = \dfrac{\partial \mathbf{H}}{\partial \boldsymbol{\rho}} \boldsymbol{C}$ can be written as

$$\begin{aligned} \frac{\partial \mathbf{H}}{\partial \boldsymbol{r}} = \frac{\partial \mathbf{H}}{\partial \boldsymbol{\rho}} \boldsymbol{C} &= \begin{bmatrix} \dfrac{\partial \|\boldsymbol{\rho}\|}{\partial \rho_u} & \dfrac{\partial \|\boldsymbol{\rho}\|}{\partial \rho_e} & \dfrac{\partial \|\boldsymbol{\rho}\|}{\partial \rho_n} \\ \dfrac{\partial az}{\partial \rho_u} & \dfrac{\partial az}{\partial \rho_e} & \dfrac{\partial az}{\partial \rho_n} \\ \dfrac{\partial el}{\partial \rho_u} & \dfrac{\partial el}{\partial \rho_e} & \dfrac{\partial el}{\partial \rho_n} \end{bmatrix} \boldsymbol{C} \\ &= \begin{bmatrix} \dfrac{\rho_u}{\|\boldsymbol{\rho}\|} & \dfrac{\rho_e}{\|\boldsymbol{\rho}\|} & \dfrac{\rho_n}{\|\boldsymbol{\rho}\|} \\ 0 & \dfrac{\rho_n}{\rho_e^2 + \rho_n^2} & \dfrac{-\rho_e}{\rho_e^2 + \rho_n^2} \\ \dfrac{\rho_e^2 + \rho_n^2}{\|\boldsymbol{\rho}\|^2 \sqrt{\rho_e^2 + \rho_n^2}} & \dfrac{-\rho_u \rho_e}{\|\boldsymbol{\rho}\|^2 \sqrt{\rho_e^2 + \rho_n^2}} & \dfrac{-\rho_u \rho_n}{\|\boldsymbol{\rho}\|^2 \sqrt{\rho_e^2 + \rho_n^2}} \end{bmatrix} \boldsymbol{C} \end{aligned} \tag{4.45}$$

4.3.4 Numerical Simulation Setup

The satellite under consideration has the following orbit parameters: $a = 6778.136\,km$, $e = 1.0 \times 10^{-5}$, $i = 51.6°$, $\omega = 30^o$, and $\Omega = 25°$. The J_2 and drag perturbation ($\boldsymbol{a}_G$ and $\boldsymbol{a}_D$ from Eq. (4.26)) are considered as the noise to the system. The location of the sensor is chosen to be the Eglin Air Force Base with 30.2316° latitude and 86.2147° west longitude. The measurement errors are assumed to be Gaussian random processes with zero means and variances of $\sigma_{range} = 25.0\ m$, $\sigma_{azimuth} = 0.015°$,

and $\sigma_{elevation} = 0.015^{\circ}$, respectively.

The true initial values of the state vector are set to be $x_0 = 4011.5713\ km$, $y_0 = 4702.6493\ km$, $z_0 = 3238.3582\ km$, $\dot{x}_0 = -5.653084\ km/s$, $\dot{y}_0 = 1.5401902\ km/s$, and $\dot{z}_0 = 4.7765408\ km/s$. For the filter, the initial values are obtained using the Herrick-Gibbs method [52], and they are $\hat{x}_0 = 3931.3399\ km$, $\hat{y}_0 = 4608.5963\ km$, $\hat{z}_0 = 3173.5911\ km$, $\hat{\dot{x}}_0 = -5.540022\ km/s$, $\hat{\dot{y}}_0 = 1.5093864\ km/s$, and $\hat{\dot{z}}_0 = 4.6810100\ km/s$.

As stated above, the acceleration due to J_2, which is approximately $10^{-5}\ km/s^2$ at the low earth orbit is considered as the noise to the system. So the process noise covariance matrix $\mathbf{Q}(t)$ is set to be $diag([0\ 0\ 0\ 10^{-10}\ 10^{-10}10^{-10}])$

The simulation of the Keplerian dynamic through the DQMOM is done using canonical unit instead of the standard SI unit [61], which means (4.28) is nondimensionalized for a better numerical stability. The initial position is used as the distance unit (DU). The velocity unit (VU) is set by $\sqrt{\mu/\mathrm{DU}}$, so the time unit (TU) is naturally equal to DU/VU. The measurement update was done in SI unit. The Keplerian equation in canonical unit has the same form as it in SI unit (the details are in Appendix B).

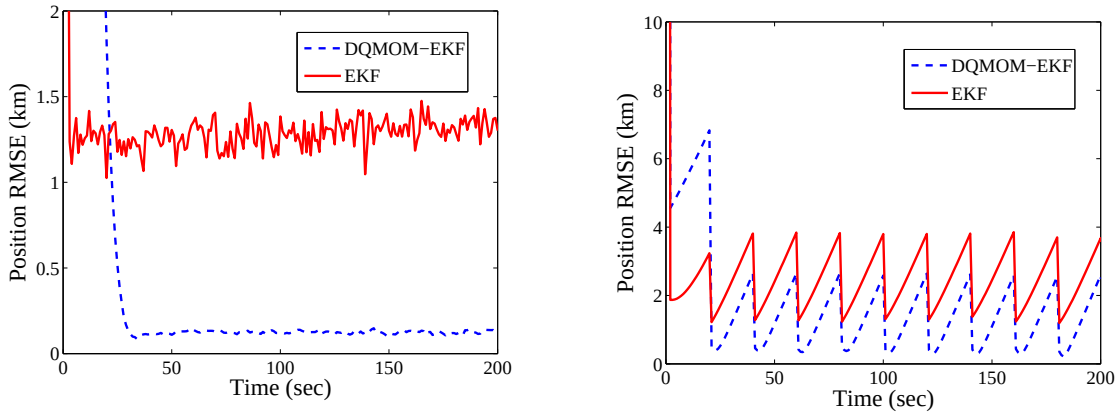
The number of the nodes used for this study was two, and the moment constraints are chosen so that the the first three moments of the PDF were preserved.

$$\begin{aligned} k1 &= [0100000 \quad 2000001] & k2 &= [0010000 \quad 0200000] \\ k3 &= [0001000 \quad 0020001] & k4 &= [0000100 \quad 0002000] \\ k5 &= [0000010 \quad 0000201] & k6 &= [0000001 \quad 0000020] \end{aligned} \tag{4.46}$$

The details regarding how to set the moment constraints can be found from Ref. [75].

4.3.5 Simulation Results

Simulations are executed with two different measurement update frequencies, 1Hz and 0.05Hz, and the result from both the DQMOM-EKF and EKF are presented. The figures are the root mean square errors (RMSE) of the position and the velocity estimation produced by Monte Carlo simulation of 30 runs. When the measurement update is frequent (1Hz), as shown in both Fig. 4.9(a) and Fig. 4.10(a), The DQMOM-EKF performs better in terms of quicker convergence in velocity estimation and better estimation accuracy in both position and velocity estimation. When the



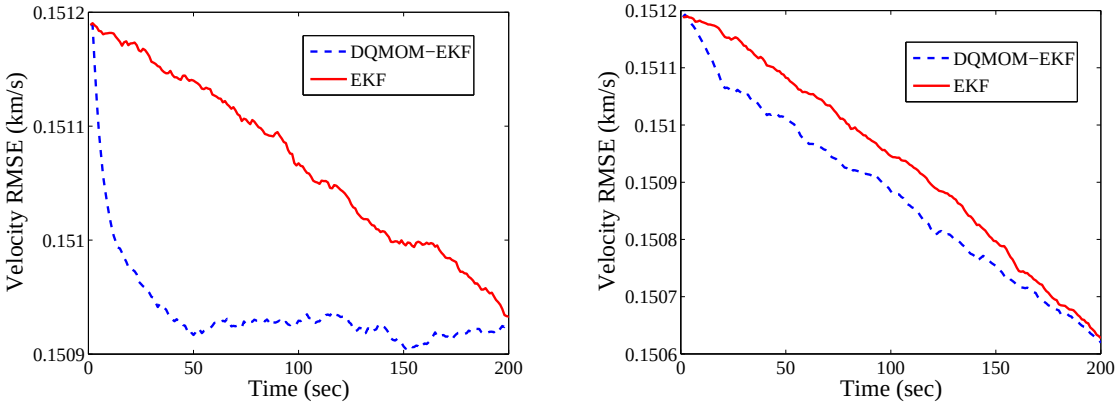
(a) Measurement update every 1 second

(b) Measurement update every 20 second

Figure 4.9: Absolute magnitude of position RMSE with different measurement update delay

time between measurement update is increased to twenty seconds (0.05Hz update frequency), the DQMOM-EKF shows quicker convergence and better estimation accuracy in both position and velocity estimation than the EKF as shown in both Fig. 4.9(b) and Fig. 4.10(b). Unlike the position RMSE curve shown in Fig.4.9(b), the velocity RMSE curve is smooth and not zigzagged. This is the result of using a fixed error covariance matrix.

Notice that it takes roughly 70 seconds in CPU time to finish the 200-second simulation for the DQMOM-EKF method, while the EKF did it in only 0.65 seconds. However, as comparing with other numerical approaches used in nonlinear filtering



(a) Measurement update every 1 second

(b) Measurement update every 20 second

Figure 4.10: Absolute magnitude of velocity RMSE with different measurement update delay

design, such as in [16, 77, 79], the computational cost is dramatically reduced and very close to the real-time estimation.

4.4 Summary of the Chapter

The ADI method is applied to a relative orbital position estimation problem. The decoupled nature of Hill's equations allowed us to solve the FPE as two low dimensional problems instead of one high dimensional problem. This helps to reduce computational cost further. The moving domain scheme that translates the domain during the numerical evaluation of FPE results in the reduction of computational time while preserving the accuracy of the estimation. The performance of the proposed method is compared with an EKF and an explicit method. The advantages of the method are: (1) as the measurement update rate decreases, the ADI based nonlinear filter methodology performs better than the EKF in terms of estimation accuracy and consistency; (2) the computational cost of the ADI is only 1/5 of the explicit method without compromise in accuracy.

The proposed hybrid filtering method is tested with a simple Bearing-only tracking problem under different measurement update rates, and the measurement update

mechanism of the EKF and UKF are used and test for their performances. Simulation results demonstrate better performances of the hybrid filter as compared with the standard Kalman filter for both frequent and sparse measurement update in terms of the estimate accuracy and convergence rate. When the measurement update rate is 1Hz it has been shown that the estimation of the proposed nonlinear filters have a better accuracy than those obtained from the EKF/UKF. With less frequent measurement updates rates (0.2 Hz) Overall the DQMOM-EKF is the best performer.

CHAPTER 5

Summary and Future Work

5.1 Summary

Since the late 1960s, the extended Kalman filter (EKF) has been used in many engineering fields that require an estimation of the state. However, the disadvantages associated with the EKF (especially for the system with high nonlinearities) have led many researchers to look for more accurate nonlinear filtering algorithms as alternatives to the EKF. Such efforts have led to algorithmic improvements in the EKF and development of new filtering methods such as the unscented Kalman filter (UKF) and the particle filter. However, these filtering methods have shown the following drawbacks

- When the system and/or measurement models are highly nonlinear, the filter can give a poor or even unstable performance as a consequence of the linearization involved in the filtering algorithm.
- The derivation of the Jacobian, which is part of the linearization process, is not a trivial task in many applications.
- Numerical evaluation of the Jacobian can be computationally intensive.
- High computational cost and memory use.

The goal of this dissertation is to address these drawbacks by solving the nonlinear filtering problem in more general framework.

In this dissertation, the existing filtering algorithms from the sequential linear least square filters to the particle filters are reviewed. For the Kalman filter, it is shown that it is the special case of the general filtering algorithm using the Fokker-Planck equation (FPE) and Bayes' rule for the linear system with the Gaussian process. The UKF is based on the unscented transformation in which a collection of sigma points are used to represent the mean and covariance of random variables undergoing nonlinear transformation efficiently, and it successfully mitigates problems involved with Jacobian calculation in the EKF. For general nonlinear and/or non-Gaussian filtering problems, the sequential Monte Carlo method is reviewed. The sequential Monte Carlo can be loosely defined as a simulation-based method that uses a Monte Carlo simulation scheme in order to solve on-line estimation and prediction problems. The flexible nature of the Monte Carlo simulations gives the particle filter better adaptability for more complex systems.

The focus of the dissertation is to find an efficient solution of the nonlinear filtering problem by improving the computational efficiency of the numerical solution of the FPE. This has been achieved by two different ways. The first one is through the finite difference methods with moving domain, and the second one is by employing the DQMOM.

To improve the finite difference method, an efficient and simple adaptive moving domain that can be applied to both the explicit and the alternating direction implicit (ADI) methods is developed to enhance the computational efficiency in solving the FPE. Unlike other moving domain schemes that adjust the domain only after each measurement update, the proposed moving domain scheme allows the modification of the domain in the middle of the numerical evaluation of the FPE. This reduced the size of the domain a lot and consequently increases in computational efficiency while preserving the accuracy. At the same time, a form of the ADI method has been derived through the implicit Euler method and the up-wind differencing scheme so

that it can be applied to many of the nonlinear dynamic problems without further derivation.

It was found that the use of the ADI method with the moving domain approach is still not fast enough for high dimensional problems. So, a more efficient numerical method, called DQMOM, is used to solve the FPE. The DQMOM converts the FPE, i.e., a partial differential equation, into a set of algebraic differential equations. However, when the update equation derived according to Bayes' formula is used together with DQMOM the "degeneracy" phenomenon, similar to the one found in a particle filter, is observed. To mitigate this problem, hybrid filtering algorithms are formed by employing the update equations from the EKF and the UKF. In addition, a fixed error covariance matrix is used for the gain calculation to prevent the filter smugness that happened in the orbit determination problem.

The proposed filtering algorithms are successfully applied to the selected nonlinear filtering problems: i) the bearing-only tracking problem, ii) the relative orbit position estimation, and iii) the orbit determination.

The ADI method with the proposed moving domain is applied to a relative orbital position estimation problem. The decoupled nature of Hill's equation allows us to solve the FPE as two low dimensional problems instead of one high dimensional problem, which further reduces the computational cost. The performance of the proposed method is compared with the EKF and the explicit method. The advantages of the method are: (1) as the measurement update rate decreases, the ADI based nonlinear filter methodology performs better than the EKF in terms of the estimation accuracy and consistency; (2) the computational cost of the ADI method is only 1/5 of the explicit method.

The proposed hybrid filtering method is tested with a simple bearing-only tracking problem under different measurement update rates. Simulation results demonstrate better performances as compared with those of the standard Kalman filter for both

frequent and sparse measurement updates. When the measurement update rate is 1Hz, it has been shown that the estimation of the proposed nonlinear filters have a better accuracy than those obtained from the EKF/UKF. With less frequent measurement update rates, i.e., 0.2 Hz, the combination of DQMOM and the update of the EKF demonstrated the best overall performance.

A nonlinear filtering algorithm utilizing the DQMOM and the EKF measurement update is used to obtain accurate and efficient orbit estimation. In addition, a fixed error covariance matrix is used for the gain calculation to prevent the filter becoming smug. A nondimensionalized system equation is used for the DQMOM to achieve a more stable propagation of the conditional PDF. Simulation results indicate that the performances of the hybrid filter based on the DQMOM and the EKF update is superior to the standard extended Kalman filter for both frequent and sparse measurement updates in terms of estimate accuracy and convergence rate. The advantages of the proposed nonlinear filtering algorithm show its potential to be suitable for efficient real-time satellite orbit determination.

5.2 Future Work

The general form of the ADI method has been developed through the implicit Euler method and the up-wind differencing scheme. The accuracy of the estimation is limited by the accuracy of the chosen numerical method. A higher accuracy, $O(\Delta t^2)$ in time, can be achieved by using D'Yaknove scheme instead of the implicit Euler method used in this dissertation which is accurate only up to $O(\Delta t)$. The up-wind scheme in this dissertation has accuracy of $O(\Delta x)$ while there exists an up-wind scheme with $O(\Delta x^2)$. However, it will make the implementation of the filter very complicated. However, by employing the numerical method with a second order accuracy the accuracy of the estimations should be improved.

The size of the adaptive moving domain does not change with respect to the size

of the PDF. So, the computation time is wasted on the part of the domain that does not contribute much on the solution if the initial PDF is much larger than the evolved one. The combination of the moving domain method in this work and the Chebyshev's inequality theory may produce better results since this will allow the method for this work to adaptively change the size of the domain.

The "degeneracy" problem presented in the filter based on the DQMOM and Bayes' rule approach was primarily resolved by constructing a hybrid filtering algorithm using the update equations from the EKF and the UKF. Even though this approach demonstrated its effectiveness, it is limiting the potential of the DQMOM. Since the cause of the phenomenon is might be that the update equation based on Bayes' rule updates only the weights, it is desirable to develop the update method based on Bayes' rule that can update not only weight but also abscissas simultaneously. Since it will mitigate the problem without borrowing the update equations from other filtering methods, the performance of the DQMOM based nonlinear filter might be better.

Bibliography

[1] Numerical evaluation of path-integral solutions to fokker-planck equations. iii. time and functionally dependent coefficients. *Phys. Rev. A*, 35:1795–1801, 1987.

[2] L. Aggoun and R. J. Elliott. *Measure Theory and Filtering*. Cambridge University Press, Cambrige, U.K., 2005.

[3] B. D. O. Anderson and J. B Moore. *Optimal Filtering*. Prentice Hall, Englewood Cliffs, NJ, 1979.

[4] Dale A. ANDERSON, John C. TANNEHILL, and Richard H. PLETCHER. *Computational Fluid mechanics and heat transfe*. Taylor & Francis, Washington, DC, 2nd edition, 1997.

[5] I. Arasaratnam, S. Haykin, and R.J. Elliott. Discrete-time nonlinear filtering algorithms using gauss hermite-quadrature. *Proceedings of the IEEE*, 95(5):953–977, May 2007.

[6] Peter J. Attar and Prakash Vedula. Direct quadrature method of moments solution of the fokker-planck equation. *Journal of Sound and Vibration*, 317(1-2):265 – 272, 2008.

[7] Peter J. Attar and Prakash Vedula. Direct quadrature method of moments solution of fokker-planck equations in aeroelasticity. *AIAA Journal*, 47(5):1219–1227, 2009.

[8] X. Lin T. Kirubarajan Y. Bar-Shalom and S. Maskell. Comparison of ekf pseudomeasurement filter and particle filter for a bearing-only tracking problem. In *Proceedings of the SPIE Conference of Signal and Data Processing of Small Targets*, pages 184–196, Orlando, FL, 2002.

[9] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation*. John Wiley & Sons, Inc., New York, NY, 2001.

[10] B. F. Spencer Jr.and L. A. Bergman. On the numerical solution of the fokker-planck equation for nonlinear stochastic systems. *Nonlinear Dynamics*, 4(4):357–372, 1993.

[11] G. J. Bierman. *Factorization Methods for Discrete Sequential Estimation*. Academic Press, New York, NY, 1977.

[12] D. Brigo, B. Hanzon, and F. LeGland. A differential geometric approach to nonlinear filtering: the projection filter. *Automatic Control, IEEE Transactions on*, 43(2):247–252, Feb 1998.

[13] R.G. Brown and P.Y. C Hwhang. *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley & Sons, Inc., New York, NY, 3rd edition, 1997.

[14] J. Carpenter, P. Clifford, and P. Fearnhead. Improved particle filter for nonlinear problems. *Radar, Sonar and Navigation, IEE Proceedings* -, 146(1):2–7, February 1999.

[15] S. Challa and Y. Bar-Shalom. Nonlinear filter design using fokker-planck-kolmogorov probability density evolutions. *IEEE Transactions on Aerospace and Electronic Systems*, 36(1):309–315, January 2000.

[16] S. Challa and Y. Bar-Shalom. Nonlinear filter design using fokker-planck-kolmogorov probability density evolutions. *IEEE Transactions on Aerospace and Electronic Systems*, 36(1):390–315, January 2000.

[17] J. L. Crassidis and J. L. Junkins. *Optimal Estimation of Dynamic Systems*. Chapman & Hall/CRC, New York, NY, 2004.

[18] F. Daum. Mesh-free adjoint methods for nonlinear filters. In *Proceedings of the SPIE Signal and Data Processing of Small Targets*, pages 109–118, San Diego, CA, 2005. SPIE.

[19] F. Daum. Nonlinear filters: Beyond the kalman filter. *IEEE Aerospace and Electronic Systems Magazine*, pages 57–69, August 2005.

[20] P.M. Djuric, J.H. Kotecha, Jianqui Zhang, Yufei Huang, T. Ghirmai, M.F. Bugallo, and J. Miguez. Particle filtering. *Signal Processing Magazine, IEEE*, 20(5):19–38, September 2003.

[21] A. Doucet, Godsill S., and C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. 10(3):197–208, 2000.

[22] Arnaud Doucet, Nando De Freitas, and Neil Gordon, editors. *Sequential Monte Carlo Methods in Practice*. Springer, New York, NY, 2001.

[23] A. N. Drozdov and M. Morillo. Solution of nonlinear fokker-planck equations. *Phys. Rev. E*, 54:931–937, 1996.

[24] A. Fokker. Die mittlerer energie rotierender elektrischer dipole im strahlungsfeld. *Annalen der Physik*, 43:810–820, 1914.

[25] Till Daniel Frank, editor. *Nonlinear Fokker-Planck Equations: Fundamentals and Applications*. Springer Series in Synergetics. Springer Berlin Heidelberg, New York, NY, 2005.

[26] K. F. Gauss. *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections, A Translation of Theoria Motus*. Dover Publications, New York, NY, 1963.

[27] A. Gelb. *Applied Optimal Estimation.* MIT Press, Cambridge, MA, 1974.

[28] A. Germani and M. Piccioni. *A Galerkin Approximation for the Zakai Equation.* Lecture Notes in Control and Information Sciences. Springer, New York, NY, 1984.

[29] Gene H. Golub and Charles F. Van Loan. *Matrix computations (3rd ed.).* Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[30] N. Gordon, D. Salmond, and A. F. M. Smith. Bayesian state estimation for tracking and guidance using the bootstrap filter. 18(6):1434–1443, 1995.

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

[32] Mohinder S. Grewal and Angus P. Andrews. *Kalman Filtering : Theory and Practice Using MATLAB.* John Wiley & Sons, Inc., New York, NY, January 2008.

[33] Ed. H. W. Sorenson. *Kalman Filtering: Theory and Application.* IEEE, Piscataway, NJ, 1985.

[34] J. M. Hammersley and K. W. Morton. Poor Man's Monte Carlo. *Journal of the Royal Statistical Society. Series B*, 16(1):23–38, 1954.

[35] Simon Haykin. *Kalman Filtering and Neural Networks.* John Wiley & Sons, Inc., New York, NY, October 2001.

[36] Y. C. Ho and R. C. K. Lee. A bayesian approach to problems in stochastic estimation and control. *IEEE Transactions on Automatic Control*, 9:333–339, October 1964.

[37] Kazufumi Ito. Approximation of the zakai equation for nonlinear filtering. *SIAM J. Control Optim.*, 34(2):620–634, 1996.

[38] A. H. Jazwinski. *Stochastic Process and Filtering Theory.* Academic Press, New York, 1970.

[39] S. Julier. A skewed approach for filtering. In *The 12th International Symposium on Aerospace Defense Sensing, Simulation and Controls*, Orlando, FL, 1998.

[40] S. Julier and J. Uhlmann. A new extension of the kalman filter to nonlinear systems. In *Int. Symp. Aerospace/Defense Sensing, Simul. and Controls, Orlando, FL*, pages 182–193, 1997.

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

[42] S.J. Julier. The scaled unscented transformation. In *American Control Conference, 2002. Proceedings of the 2002*, volume 6, pages 4555–4559 vol.6, 2002.

[43] S.J. Julier and J.K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, March 2004.

[44] R.E Kalman. A new approach to linear filtering and prediction problems. *ASME–Journal of Basic Engineering*, 82D:34–45, March 1960.

[45] R.E Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *ASME–Journal of Basic Engineering*, pages 95–108, March 1961.

[46] Genshiro Kitagawa. Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 5(1):1–25, 1996.

[47] Augustine Kong, Jun S Liu, and Wing Hung Wong. Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*, 89:278–288, 1994.

[48] S. Musick J. Greenswald C. Kreucher and K. Kastella. Comparison of particle method and finite difference nonlinear filters for low snr target tracking. In *4th International Conference on Information Fusion*, 2001.

[49] S. Musick J. Greenswald C. Kreucher and K. Kastella. Comparison of particle method and finite difference nonlinear filters for low snr target tracking. In *The 2001 Defense Applications of Signal Processing Workshop*, 2001.

[50] R. S. Langley. A finite element method for the statistics of non-linear random vibration. *Journal of Sound and Vibration*, 101:41–54, 1985.

[51] H. P. Langtangen. A general numerical solution method for fokker-planck equations with applications to structural reliability. 6:33–48, 1991.

[52] D Lee and K. Alfriend. Sigma point filtering for sequential orbit estimation and prediction. *Journal of Spacecraft and Rockets*, 44:388–398, 2007.

[53] F. L. Lewis. *Optimal Estimation with an Introduction to Stochastic Control Theory*. John Wiley & Sons, Inc., New York, NY, 1986.

[54] Jun S. Liu and Rong Chen. Sequential monte carlo methods for dynamic systems. *Journal of the American Statistical Association*, 93:1032–1044, 1998.

[55] Sergey Lototsky, Remijigus Mikulevicius, and Boris L. Rozovskii. Nonlinear filtering revisited: A spectral approach. *SIAM Journal on Control and Optimization*, 35:435–461, 1997.

[56] A. MacCormick, J.and Blake. A probabilistic exclusion for tracking multiple objects. 39(1):57–71, 2000.

[57] Daniele L. Marchisio and Rodney O. Fox. Solution of population balance equations using the direct quadrature method of moments. *Journal of Aerosol Science*, 36(1):43 – 73, 2005.

[58] S. Maskell and N. Gordon. A tutorial on particle filters for on-line nonlinear/non-gaussian bayesian tracking. In *Target Tracking: Algorithms and Applications (Ref. No. 2001/174), IEE*, volume Workshop, pages 2/1–2/15 vol.2, October 2001.

[59] P. S. Maybeck. *Stochastic Models, Estimation and Control. Volume I.* New York, NY, 1979.

[60] F. Michael and M. D Johnson. Financial market dynamics. *Physica A: Statistical Mechanics and its Applications*, 320:525–534, March 2003.

[61] Roger R. Bate Donald D. Mueller and Jerry E. White. *Fundamentals of aerodynamics*, chapter 2. Dover, New York, NY., 1971.

[62] M. Planck. Über einen satz der statistischen dynamik und seine erweiterung in der quantentheorie. *Sitzungsber. Preuss.Akad. Wissens*, pages 324–341, 1917.

[63] Sun J. Q. and Hsu C. S. The generalized cell mapping method in nonlinear random vibration based upon short-time gaussian approximation. *ASME Journal of Applied Mechanic*, 57:1018–1025, 1990.

[64] B. Ripley. *Stochastic Simulation.* John Wiley & Sons, Inc., New York, NY, 1987.

[65] Branko Ristic, Sanjeev Arulampalam, and Neil Gordon. *Beyond the Kalman Filter: Particle Filters for Tracking Applications.* Artech House, Boston, MA, 2004.

[66] A. F. M. Smith and A. E. Gelfand. Bayesian statistics without tears: A sampling-resampling perspective. *The American Statistician*, 46(2):84–88, 1992.

[67] H. W Sorenson. On the development of practical nonlinear filters. *Information Sciences*, 7:253–270, 1974.

[68] S. Sadhua S. Mondalb M. Srinivasana and T. K. Ghoshal. Sigma point kalman filter for bearing only tracking. *Signal Processing*, 86(12):3769–3777, December 2006.

[69] R. F. Stengle. *Optimal Control and Estimation.* Dover Publications, New York, NY, 1994.

[70] J. R. Strikwerda. *Finite difference schemes and partial differential equations.* New York Chapman and Hall, New York, NY, 1989.

[71] H. Tanizaki. *Nonlinear Filters: Estimation and Applications.* Springer, New York, NY, 1996.

[72] David A. Vallado. *Fundamentals of astrodynamics and applications*, chapter 6. Microcosm Press, New York, NY, 2nd edition, 2001.

[73] E. A. Wan and R. Van Der Merwe. The unscented kalman filter for nonlinear estimation. In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pages 153–158, 2000.

[74] Yuanxin Wu, Dewen Hu, Meiping Wu, and Xiaoping Hu. Unscented kalman filtering for additive noise case: augmented versus nonaugmented. *Signal Processing Letters, IEEE*, 12(5):357–360, May 2005.

[75] Y. Xu and P. Vedula. A quadrature based method of moments for nonlinear filtering. *Automatica*, 45(5):1291–1298, May 2009.

[76] Y. Xu and J. Yoon. A direct quadrature approach for nonlinear filtering. In *The 2009 American Control Conference*, St. Louis, Missouri, June 2009. ACC.

[77] J. Yoon and Y. Xu. Alternating directional implicit method enhanced nonlinear filtering for relative orbit estimation. *Journal of Aerospace Engineering to be published*.

[78] J. Yoon and Y. Xu. Relative position estimation using fokker-planck and bayes' equations. In *AIAA Guidance, Navigation and Control Conference and Exhibit*, Hilton Head, South Carolina, August 2007. AIAA.

[79] J. Yoon, Y. Xu, and P. Vedula. Enhanced direct quadrature based method of moments for nonlinear filtering. In *AIAA Guidance, Navigation and Control Conference and Exhibit*, Chicago, Illinois, August 2009. AIAA.

[80] A Zatezalo. Tracking and detection for the target state model. Master's thesis, University of Minnesota, Minneapolis, MN, March 1997.

[81] D. S. Zhang, G. W. Wei, D. J. Kouri, and D. K. Hoffman. Numerical method for the nonlinear fokker-planck equation. 56:1197–1206, 1997.

APPENDIX A

Calculation of the Local Sidereal Time

The local sidereal time is the angle between the x-axis of the geocentric-equatorial coordinate (vernal equinox direction) and the longitude of the sensor location on the Earth. It is the sum of the Greenwich sidereal time which is the angle between the Greenwich Meridian and the x-axis of the geocentric-equatorial coordinate and the geographical longitude of the location.

$$\theta = \theta_{GMT} + \phi \tag{A.1}$$

where θ_{GMT} is the Greenwich sidereal time and ϕ is the geographical longitude of the sensor location.

There are two different ways to calculate the local sidereal time (LST). The first method requires knowing the Greenwich sidereal time (GST) at the beginning of the particular year (θ_{GMT_0}) in UT_1 that can be found from Astronomical Almanac [61, 72] where UT_1 is the universal time in solar seconds. Once the θ_{GMT_0} of particular year (January 1, 0 hour, 0 minute, 0 second) is found, the LST can be found by

$$\theta = \theta_{GMT_0} + 1.002737909350795 \times 2\pi \times D + \phi \tag{A.2}$$

where D is the total elapsed time in solar days from θ_{GMT_0}, e.g. D for Jan 31 is 30. θ found from this equation is in radian.

The second method requires calculating θ_{GMT_0} and is a little more complex. The first step is to find the Julian date (JD) of the time in UT_1. JD is the Julian day

numbers for the date of interest for the epoch J2000 that also can be found from the Astronomical Almanac and is simply the integer part of the JD, i.e., the JD at 0 h 0 min 0 s of the day. It is computed by

$$JD = (367 \times yr) - INT\left\{\frac{7\left\{yr + \text{INT}\left(\frac{mo+9}{12}\right)\right\}}{4}\right\} + \text{INT}\left(\frac{275mo}{9}\right) + d + 1721013.5 + \frac{\frac{\left(\frac{s}{60}+min\right)}{60} + h}{24} \tag{A.3}$$

where yr, mo, d, h, min, and s are full digit year, month, day, hour, minute, second accordingly and INT means real truncation. This equation is valid for the time period from March 1, 1900 to February 28, 2100. The second step is finding the Julian century of the Julian date by

$$T = \frac{JD - 2451545.0}{36525} \tag{A.4}$$

With this the Greenwich sidereal time is

$$\theta_{GMT} = 67310.548.31^s + \left(876600^h\frac{3600^s}{1^h} + 8640184.812866^s\right)T + 0.093104T^2 - 6.2 \times 10^{-6}T^3 \tag{A.5}$$

The superscripts, s and h denote seconds and hours in angular measure not in time. θ_{GMT} from this calculation is in second which is 240^{th} of degree. After converting it in degree the LST is $\theta = \theta_{GMT} + \phi$ (degree) where ϕ is the geographical longitude that east is taken to be positive.

APPENDIX B

Nondimensionalization of the Kepler Equation

The simulation of the Keplerian dynamic through the DQMOM is done using canonical unit instead of the standard SI unit [61], which means Eq. (B.1) is nondimensionalized for a better numerical stability.

$$\ddot{\boldsymbol{r}} = \frac{d\boldsymbol{v}}{dt} = -\frac{\mu}{r^3}\boldsymbol{r} \tag{B.1}$$

The radius of the initial position is used as the distance unit $\overline{r} = r_0$, and the time unit is naturally equal to $\overline{t} = \overline{r}/\overline{v}$ where $\overline{v}$ is the velocity unit defined by $\overline{v} = \sqrt{\mu/\overline{r}}$. With these canonical units, the position, the velocity, the time, and the gravitational parameter can be redefined as

$$\boldsymbol{r} = \hat{\boldsymbol{r}}\overline{r} \qquad \boldsymbol{v} = \hat{\boldsymbol{v}}\overline{v} \qquad t = \hat{t}\overline{t} = \hat{t}\frac{\overline{r}}{\overline{v}} \qquad \mu = \hat{\mu}\overline{v}^2\overline{r} \tag{B.2}$$

where $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{v}}$, $\hat{t}$ and $\hat{\mu}$ are the nondimesional position, the velocity, the time and the gravitational parameter, respectively.

The nondimensional form of the Eq. (B.1) is

$$\ddot{\hat{\boldsymbol{r}}} = \frac{d\hat{\boldsymbol{v}}}{d\hat{t}} = \frac{d}{dt\frac{\overline{v}}{\overline{r}}}\left(\frac{\boldsymbol{v}}{\overline{v}}\right) = \frac{\overline{r}}{\overline{v}}\left[\frac{d\boldsymbol{v}}{dt}\left(\frac{1}{\overline{v}}\right) + \frac{d}{dt}\left(\frac{1}{\overline{v}}\right)\boldsymbol{v}\right] \tag{B.3}$$

Since $\overline{v}$ is constant, $\dfrac{d}{dt}\left(\dfrac{1}{\overline{v}}\right) = 0$, it becomes

$$\ddot{\hat{\boldsymbol{r}}} = \frac{d\boldsymbol{v}}{dt}\left(\frac{\overline{r}}{\overline{v}^2}\right) = -\frac{\mu}{r^3}\boldsymbol{r}\left(\frac{\overline{r}}{\overline{v}^2}\right) \tag{B.4}$$

By replacing the gravitational parameter μ and $\boldsymbol{r}$, using above definition, it can be rewritten as

$$\ddot{\hat{\boldsymbol{r}}} = \frac{d\boldsymbol{v}}{dt}\left(\frac{\overline{r}}{\overline{v}^2}\right) = -\frac{\hat{\mu}\overline{v}^2\overline{r}}{r^3}\hat{\boldsymbol{r}}\overline{r}\left(\frac{\overline{r}}{\overline{v}^2}\right) = -\frac{\hat{\mu}}{r^3}\hat{\boldsymbol{r}}\overline{r}^3 \tag{B.5}$$

Finally, using $r^3 = \hat{r}^3\overline{r}^3$, Eq. (B.1) can be written in complete nondimensional form as

$$\ddot{\hat{\boldsymbol{r}}} = -\frac{\hat{\mu}}{\hat{r}^3}\hat{\boldsymbol{r}} \tag{B.6}$$

The Keplerian equation in canonical unit has the same form as it in SI unit.