NEW ALGORITHMS FOR UNCERTAINTY QUANTIFICATION AND NONLINEAR ESTIMATION OF STOCHASTIC DYNAMICAL SYSTEMS

A Dissertation

by

PARIKSHIT DUTTA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2011

Major Subject: Aerospace Engineering

NEW ALGORITHMS FOR UNCERTAINTY QUANTIFICATION AND NONLINEAR ESTIMATION OF STOCHASTIC DYNAMICAL SYSTEMS

A Dissertation

by

PARIKSHIT DUTTA

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Raktim Bhattacharya
Committee Members,	Suman Chakravorty
	John L. Junkins
	Aniruddha Datta
Head of Department,	Dimitris C. Lagoudas

August 2011

Major Subject: Aerospace Engineering

ABSTRACT

New Algorithms for Uncertainty Quantification and Nonlinear Estimation of Stochastic Dynamical Systems. (August 2011)

Parikshit Dutta, B.Tech.; M.Tech., Indian Institute of Technology, Kharagpur

Chair of Advisory Committee: Dr. Raktim Bhattacharya

Recently there has been growing interest to characterize and reduce uncertainty in stochastic dynamical systems. This drive arises out of need to manage uncertainty in complex, high dimensional physical systems. Traditional techniques of uncertainty quantification (UQ) use local linearization of dynamics and assumes Gaussian probability evolution. But several difficulties arise when these UQ models are applied to real world problems, which, generally are nonlinear in nature. Hence, to improve performance, robust algorithms, which can work efficiently in a nonlinear non-Gaussian setting are desired.

The main focus of this dissertation is to develop UQ algorithms for nonlinear systems, where uncertainty evolves in a non-Gaussian manner. The algorithms developed are then applied to state estimation of real-world systems. The first part of the dissertation focuses on using polynomial chaos (PC) for uncertainty propagation, and then achieving the estimation task by the use of higher order moment updates and Bayes rule. The second part mainly deals with Frobenius-Perron (FP) operator theory, how it can be used to propagate uncertainty in dynamical systems, and then using it to estimate states by the use of Bayesian update. Finally, a method to represent the process noise in a stochastic dynamical system using a finite term Karhunen-Loève (KL) expansion is proposed. The uncertainty in the resulting approximated system is propagated using FP operator.

The performance of the PC based estimation algorithms were compared with extended Kalman filter (EKF) and unscented Kalman filter (UKF), and the FP operator based techniques were compared with particle filters, when applied to a duffing oscillator system and hypersonic reentry of a vehicle in the atmosphere of Mars. It was found that the accuracy of the PC based estimators is higher than EKF or UKF and the FP operator based estimators were computationally superior to the particle filtering algorithms.

To my family and my friends

ACKNOWLEDGMENTS

I would like to give special thanks to Dr. Raktim Bhattacharya who has been a mentor and a friend. His creativity and drive have been essential to the completion of this work. Secondly, I would like to thank Dr. Suman Chakravorty for his guidance during my time as graduate student and clarifying several important concepts. I would also like to thank Dr. John L. Junkins for inspiring me, and helping me build the fundamental understanding of estimation theory.

My appreciation also goes to all my friends who have been very supportive and have helped me to maintain my sanity while completing this work. I would especially like to thank Abhishek, Baljeet, Avinash and Ramesh with whom I have worked and shared several good times at the CISAR laboratory. Their support and inspiration were of immense help towards completion of my work. I would also like to thank several other friends for always being there to share both laughter and hardship.

Last, but not the least I would like to thank my family especially my parents for everything they have done for me, and my uncle who helped me develop my interest in science.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

A. Background

Uncertainty quantification in stochastic dynamical systems is a challenging field that has received attention for over a century. Understanding the impact of uncertainty in complex physical system has become a primary topic of research over the years. Problems due to uncertainty may arise in myriad areas of application ranging from robotics (e.g. path planning in an uncertain environment) and astrodynamics (eg. trajectory estimation of 99942 Aphophis) to structural engineering (e.g. excitation of a building caused by siesmic events) and petroleum engineering (eg. study of flow of oil through a reservoir with uncertain porosity). A large class of such problems deal with uncertainty in physical model and system parameters. Estimation of parameters in this scenario is typically a hard problem due to lack of frequent measurements and underlying nonlinearities in system dynamics. Thus the evolution of uncertainty, which can be non Gaussian, needs to be predicted over longer intervals of time. These issues undermine the validity of the classical linear Gaussian theory. Sequential estimation algorithms, based on Monte Carlo (MC) simulations are most commonly used in such cases. However for systems having three or more dimensions, MC based techniques may be computationally expensive as ensemble size required to guarantee convergence, increases exponentially with number of states. Hence a nonlinear estimation algorithm, superior to the existing methods in terms of convergence of errors and computational complexity is desired.

Estimation of states and parameters for dynamical systems in general, are gen-

---

The journal model is *IEEE Transactions on Automatic Control.*

erally performed in the Bayesian framework, where uncertainty is represented as probability density functions (PDF). For linear Gaussian systems, it is possible to get exact analytical expressions for evolving sequence of moments, which characterizes the PDF completely. This method is widely know as Kalman filter [1]. For nonlinear systems exhibiting Gaussian behavior, the system is linearized locally, about the current mean, and the covariance is propagated using the approximated linear dynamics. This method is used in extended Kalman filters (EKF) [2]. It is well known that this approach performs poorly when the nonlinearities are high, resulting in an unstable estimator [3, 4, 5, 6]. However, the error in mean and covariance can be reduced if the uncertainty is propagated, using the nonlinear dynamics, for a minimal set of sample points, called sigma points. The PDF of the states, characterized by sigma points, capture the posterior mean and covariance accurately to the third order (Taylor series expansion) for any nonlinearity with Gaussian behavior. This technique has resulted in the unscented Kalman filter (UKF) [7]. The aforementioned filters are based on the premise of Gaussian PDF evolution. If the sensor updates are frequent then EKF and UKF may yield satisfactory results. However, for nonlinear systems, if the sensor updates are slow, these filters result in inaccurate estimates [8].

Recently, simulation-based sequential filtering methods, using Monte Carlo simulations, have been developed to tackle nonlinear system with non-Gaussian uncertainty [9, 10]. Monte Carlo methods involve representing the PDF of the states using a finite number of samples. The filtering task is obtained by recursively generating properly weighted samples of the state variable using importance sampling [11]. These filters, based on sequential MC methods are known as Monte Carlo filters [12]. Amongst them, the most widely used is the particle filter [13, 14, 15, 16]. Here ensemble members or *particles* are propagated using the nonlinear system dynamics. These particles with proper weights, determined from the measurements, are used

to obtain the state estimate. However, particle filters require a large number of ensembles for convergence, leading to higher computational costs [17]. This problem is tackled through *resampling* [13, 18, 8]. Particle filters with resampling technique are commonly known as bootstrap filters [8]. It has been observed that bootstrap filters introduce other problems like loss of diversity amongst particles [15], if the resampling is not performed correctly. Recently developed techniques have combined importance sampling and Markov-Chain-Monte Carlo (MCMC) methods to generate samples to get better estimates of states and parameters [19]. Several other methods, like regularized particle filter [20], and filters involving MCMC move step [21], have been developed to improve sample diversity. At the same time, even with resampling, due to the simulation based nature of these filters, the ensemble size scales exponentially with state dimension for large problems [22]. To circumvent this problem, particle filters based on Rao-Blackwellization have been developed to partially solve the estimation problem analytically [23]. However, its application is limited to systems where the required partition of the state space is possible. An excellent comparison of the various nonlinear filtering algorithms is available in ref. [24].

Nonlinear estimation algorithms based on polynomial chaos theory [25] and Frobenius-Perron operator [26] has been proposed in this work. Polynomial chaos (PC) is used to approximate any random process as linear combination of orthogonal basis functions. The advantage of using PC is that an alternate deterministic dynamical system can be created from the stochastic system, which is then used to propagate uncertainty. Polynomial chaos was first introduced by Wiener [27] where Hermite polynomials were used to model stochastic processes with Gaussian random variables. According to Cameron and Martin [28], such an expansion converges in the $\mathcal{L}_2$ sense for any arbitrary stochastic process with finite second moment. This applies to most physical systems. Xiu *et al.* [29] generalized the result of Cameron-Martin

to various continuous and discrete distributions using orthogonal polynomials from the so-called Askey-scheme [30] and demonstrated $\mathcal{L}_2$ convergence in the corresponding Hilbert functional space. This is popularly known as the generalized polynomial chaos (gPC) framework. The gPC framework has been applied to various applications including stochastic fluid dynamics [31, 32], stochastic finite elements [25], and solid mechanics [33, 34]. In the context of nonlinear estimation, polynomial chaos has been applied by Blanchard *et al.* [35, 36], where uncertainty prediction was computed using gPC theory for nonlinear dynamical systems, and estimation was performed using linear output equations and classical Kalman filtering theory. It has been shown that PC is computationally more efficient than Monte Carlo simulations [29]. Hence, it is expected that, the estimation algorithm presented here will be computationally more efficient than particle filters. However, such an analysis has not been performed, and is a subject of our future work. Here we have applied gPC theory to estimate states of a Duffing oscillator and hypersonic vehicle reentering Mars' atmosphere.

The Frobenius-Perron operator determines the time evolution of probability density function (PDF) through a system, and as shown later is computationally efficient than particle filters. The Frobenius-Perron operator has been used in the physics community to study evolution of uncertainty in dynamical systems [26]. In continuous time, the Frobenius-Perron operator is defined by the Liouville equation [37], which is the Fokker-Planck equation [38] without the diffusion term. It has been shown that the Frobenius-Perron operator or the Liouville equation, predicts evolution of uncertainty in a more computationally efficient manner than Monte Carlo [39]. Based on this fact, we can expect a nonlinear filtering algorithm in this framework to be computationally more efficient than particle filters. However, it is important to note that the Frobenius-Perron operator only addresses parametric uncertainty. Use of Liouville equation to develop a nonlinear filtering algorithm was first presented

by Daum *et al.* [40], where the process of the filtering algorithm has been outlined. In this work we have applied Frobenius-Perron operator theory to a state estimation problem arising in hypersonic flights and perform a direct comparison with particle filters.

The above mentioned uncertainty propagation methods are applicable only when the dynamical system has initial state or parametric uncertainty. In presence of process noise, the evolution of densities are given by the Kramers-Moyal expansion [41]. This is a partial differential equation which characterize PDF propagation in any nonlinear system having process noise. A special case arises when we limit ourselves to additive Gaussian white noise as stochastic forcing. Then, the first two terms of the Kramers-Moyal expansion is sufficient to describe the evolving densities. This is referred to as the Fokker-Planck equation or Kolmogorov forward equation [42]. There are several methods, by which we can approximately determine the solution of the Fokker-Plank equation. A brief treatise of the most popular methods can be found in the book by Riskin [38]. Several methods which deal with numerical solutions of the Fokker-Planck equation have been developed over the years [43, 44, 45]. The numerical algorithms, intend to solve the Fokker-Planck equation using grid based methods like FEM, to evaluate the densities in a structured grid [46, 47, 48], or by using meshfree methods, by evaluating densities at randomly selected points to get the final PDF [49, 50]. Several researchers have used Monte Carlo based techniques to get an approximation of the solution of Fokker-Planck equation using finite number of samples [51, 50]. Another popular technique of solving Fokker-Planck equation involves approximating the PDF as linear combination of known functions. Researchers have often used known densities as the basis functions for approximation. This method is popularly known as kernel density estimation [52]. There are several techniques which solve the Fokker-Planck equation using this method [53, 54]. One

can also use known functions to get a finite dimensional approximation of the operator generated by the Fokker Planck equation. This method is useful as it converts the Fokker Planck partial differential equation into an approximate ordinary differential equation. Kumar *et al.* presents a method of solving the Fokker Planck equation using this technique [49, 55]. It has been observed that most of the solution methodologies perform poorly when the dimensionality of the state space involves is high [56]. This has been proved for grid based method as the complexity in solving the problem increases exponentially with dimension [57]. The problem is partially resolved by the use of sparse grids where a structured grid is used to evaluate the solution at lesser number of points than the grid based methods [58]. But even with sparse grids, accuracy of solution become worse with increase in dimensions [59]. For methods using approximating basis functions, finding the correct basis for evolution is challenging when one has to deal with high dimensional problem [59]. Hence most of the solution methods of the Fokker Planck equation suffer from the curse of dimensionality [60].

In this dissertation, we use a methodology based on Frobenius-Perron operator theory and Karhunen Loève expansion, to determine the sequence of evolving densities in a stochastic dynamical system. Karhunen Loève (KL) expansion has been developed independently by researchers to represent a random process as linear combination of orthogonal functions [61, 62]. KL expansion, expands any random process as homogeneous products of functions of deterministic and stochastic variables. It is widely used in physics and fluid mechanics to represent noise in a Langevin equation and turbulence models [63, 64]. In the context of dynamical systems, it has primarily been used in model reduction and data analysis of complex high dimensional systems [65, 66, 67]. KL expansion has also found applications in the areas of nonlinear vibrations [68], wavelet analysis [69, 70], and signal processing [71]. However, its application to problems involving uncertainty propagation in dynamical systems

has been limited. Here we use a methodology where the process noise in a system is represented as a KL expansion of the underlying random process, and then use Frobenius-Perron operator to propagate uncertainty. We have applied the resulting uncertainty propagation method, to an estimation problem where we estimate states of a hypersonic reentry vehicle. The results have been compared with particle filtering methods.

B. Contribution of This Dissertation

In this dissertation we mainly focus on developing new, efficient algorithms for uncertainty quantification of dynamical systems and apply them to state estimation problems. In particular we assume that the uncertainty in the system dynamics is dependent on random variable, governed by a known PDF. Throughout the dissertation our focus is representing the PDF as a continuous function of the underlying random variable. Although it is possible to extend these results to discrete distributions, these have not been treated.

The main contribution of this dissertation lies in the application of the proposed estimation algorithms to real-world problems. The problem that we focus on here is hypersonic reentry of a vehicle in the atmosphere of Mars. Entry, descent, landing of a hypersonic vehicle on the surface of Mars is a topic of research receiving much attention in recent years. The expected mass of the next Mars science mission laboratory is approximately 2800 kilograms at entry, which is required to land within few kilometers of robotic test sites. The requirement of high accuracy when landing in proximity of the target region is a key challenge of high mass entry. It is therefore necessary to estimate states and parameters of the reentry vehicle when uncertainties are present in initial conditions. High nonlinearity of reentry dynamics, coupled with

lack of frequent sensor updates make the estimation problem difficult to solve. In the subsequent chapters, we develop algorithms to effectively quantify uncertainty for the reentry vehicle and to estimate states of the reentry vehicle.

In Chapter II, we discuss some of the commonly used state estimation methods. We first introduce Kalman filter, which is optimal for linear Gaussian systems. Next we discuss some of the suboptimal algorithms for estimation of nonlinear systems. Finally we show, through an example how these estimation algorithms perform when applied to a nonlinear system.

In Chapter III, we introduce polynomial chaos (PC) and develop two relevant estimation algorithms; one that uses higher order moment updates and other using Bayesian update. We apply the proposed estimation methodologies to estimate states of a Duffing oscillator and eventually apply them to state estimation of a hypersonic reentry vehicle. We compare our results with estimators based on EKF and UKF.

Chapter IV deals with, uncertainty propagation using Frobenius-Perron (FP) operator theory. We first develop the methodology of uncertainty propagation, and then we apply the estimation scheme to hypersonic reentry vehicle and compare our results with particle filters.

In Chapter V, we propose a methodology for uncertainty quantification when the dynamical system has process noise in it. We use Karhunen Loève (KL) expansion to represent process noise and then use Frobenius-Perron operator to propagate uncertainty. We show how the proposed methodology can be used to estimate states and parameters of a stochastic dynamical system. We apply the methodology to hypersonic reentry problem and compare the results with particle filter.

Finally in Chapter VI we summarize our conclusions and highlight some future directions of research.

CHAPTER II

SEQUENTIAL STATE ESTIMATION METHODS

State estimation methods is a topic of research that has gained popularity over the years. The development of estimation methods was pioneered by Gauss, in 18^{th} century, when he proposed the *method of least squares* [72]. In the sequential estimation setting, Gaussian least squares method used to reduce the estimation error sequentially with each observations, by incrementally correcting the measurements. This is known as Gaussian least square differential correction (GLSDC) [73]. For nonlinear systems, the dynamics can be linearized about the current estimate, which can then be used for state estimation purposes. This method is known as nonlinear least squares [73]. In a probabilistic setting, this method in turn minimizes the variance of the state estimate from the true value, and hence is called the *minimum variance estimator* [74]. However, Gauss did anticipate the need of the most probable estimate of the state rather than one which minimizes the variance [75]. This was first introduced by R. A. Fisher, in 1912, as *the maximum likelihood estimator* [76]. It is interesting to note that if the states follow a Gaussian distribution the minimum variance and the maximum likelihood estimates are the same.

The first concepts of estimating states of a system, as a consequence of a filtering problem was proposed independently by Wiener and Kolmogorov, which became popularly known by the name of Wiener-Kolmogorov filter [77]. In this framework, the objective was to filter out noise from a signal by minimizing the mean square error. The filter was formulated both for continuous and discrete observations, which made it different from the least squares technique, where observations arrived in discrete time intervals. All these contributions were significant towards the development of Kalman filter which will be discussed in detail, in the subsequent sections [1].

State estimation of dynamical systems is generally done in two steps: a forward step, where the PDF of the states is propagated forward in time, to get the *prior* PDF; and an inverse step, where the prior is updated based on observations to get the *posterior* PDF. The forward step, generally reduces to propagation of uncertainty through the stochastic dynamical system, and the inverse step reduces to, using Bayesian inference [78]. The state estimate is obtained as a result of applying a desired optimization criterion on the posterior PDF.

In this chapter, we will cover, methodologies for state estimation of dynamical systems. We will begin with a brief overview of the techniques that are popularly used and then we will describe in brief the methodologies that has been proposed in this framework. Finally we would present an example of application of the state estimation methods discussed in this chapter.

A. Sequential State Estimation for Linear Systems

In this section, we propose a methodology where the dynamical system in question is a linear system. We consider systems with dynamics and measurements described by following sets of equations,

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + G(t)w(t) \tag{2.1a}$$

$$\tilde{y}(t) = H(t)x(t) + v(t) \tag{2.1b}$$

where $x \in \mathbb{R}^n$ are states, $u \in R^m$ are controls and $\tilde{y} \in \mathbb{R}^p$ are observations. $A \in \mathbb{R}^{n\times n}$ is the state transition matrix, $B \in \mathbb{R}^{n\times m}$ is the input coefficient matrix, and $H \in \mathbb{R}^{p\times n}$ is the matrix relating states to output. $w \in \mathbb{R}^q$ and $v \in \mathbb{R}^p$ are zero mean Gaussian white noise processes, and $G \in \mathbb{R}^{n\times q}$ is the process noise coefficient matrix.

We will now describe the formulation of linear state estimation algorithms. We

will consider only *continuous-discrete* cases i.e. the stochastic dynamical system is propagated continuously and the measurements arrive at discrete time intervals.

1. Kalman Filter

Kalman filter is a sequential state estimation method which gives us exact sequence of evolving densities for a linear systems with Gaussian uncertainties. It was developed by R.E. Kalman in 1960 for discrete systems [1]. The continuous time version of this filter is called the Kalman-Bucy filter which was developed in 1961 [79]. Kalman filter gives us optimal state estimate for a linear system which takes Gaussian densities to initially, and the density of the states remains Gaussian throughout the propagation time. It postulates a dynamical equation for covariance and mean, for the underlying Gaussian density function. Thus, the PDF of the states can be completely characterized evolving sequence of moments.

Given the system in Eqn. (2.1a) and Eqn. (2.1b), we assume that the process noise $w(t)$ and the measurement noise $v(t)$ are uncorrelated. Moreover in a continuous-discrete formulation, the measurement equation is assumed to be discrete. Hence Eqn. (2.1a) and Eqn. (2.1b) in this case is modified to,

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + G(t)w(t) \tag{2.2a}$$

$$\tilde{y}_k = H_k x_k + v_k \tag{2.2b}$$

The subscript k represents the time instant t_k when the measurement is available. Also $\mathbb{E}[v_k] = 0, \forall k \in \mathbb{N}$, $\mathbb{E}[v_k v_j] = R\,\delta_{kj}$, and $\mathbb{E}[w(t)] = 0, \forall t \in \mathbb{R}^+$, $\mathbb{E}[w(t_1)w(t_2)] = Q\,\delta(t_1 - t_2)$. We will assume that R and Q remain constant.

In a continuous-discrete Kalman filter, the mean $\mu_{k|k}$ and the covariance $P_{k|k}$ are propagated forward in time from current time step, t_k to the next step t_{k+1}, when the measurements are available, to get the prior mean $\mu_{k+1|k}$, and covariance $P_{k+1|k}$. The

prior mean and the covariance are then updated using the Kalman gain K to get the posterior mean $\mu_{k+1|k+1}$ and covariance $P_{k+1|k+1}$, which is obtained by minimizing the error covariance. It is interesting to note that regardless of the optimization criterion used the optimal estimate for a Kalman filter is the mean of the posterior PDF, hence $\hat{x}_{k+1} = \mu_{k+1|k+1}$.

The initial state estimate and covariance are the mean and the covariance of the initial PDF, i.e. $\hat{x}_0 = \mathbb{E}[x(t_0)]$ and $P_{0|0} = \mathbb{E}[x^2(t_0)] - \mathbb{E}[x(t_0)]^2$. The forward propagation step, for a Kalman filter essentially consists of two equations for mean and covariance propagation, which are given by,

$$\dot{\mu}(t) = A(t)\mu(t) + B(t)u(t), \text{ with } \mu(t_k) = \mu_{k|k} \tag{2.3a}$$

$$\dot{P}(t) = A(t)P(t) + P(t)A(t)^T + Q, \text{ with } P(t_k) = P_{k|k} \tag{2.3b}$$

The update step consists of solving an optimization problem to get the Kalman gain K_{k+1} at step t_{k+1} and subsequently obtaining the state estimate and the covariance of the posterior PDF. The update equations are given by,

$$K_{k+1} = P_{k+1|k}H_{k+1}^T\left(P_{k+1}P_{k+1|k}H_{k+1}^T + R\right)^{-1} \tag{2.4a}$$

$$\mu_{k+1|k+1} = \mu_{k+1|k} + K_{k+1}(\tilde{y}_{k+1} - H_{k+1}\mu_{k+1|k}) \tag{2.4b}$$

$$P_{k+1|k+1} = (\mathbf{I} - K_{k+1}H_{k+1})P_{k+1|k} \tag{2.4c}$$

The state estimate is given by the posterior mean i.e. $\hat{x}_{k+1} = \mu_{k+1|k+1}$.

2. Linear Non-Gaussian Filter

In this section we briefly describe methodology of estimation when the given PDF of initial states is not a Gaussian distribution. We still deal with linear system hence

the PDF of states undergoes only linear transformation while it evolves. Hence then structure of the PDF is conserved, but due to linear transformation the parameters by which it is represented changes.

We do not have an optimal estimator for such systems. But we can design a filter which is suboptimal by approximating the initial PDF using Gaussian Mixture Models (GMM) [80, 81]. In this framework we represent any PDF as linear combination of Gaussian PDFs. This method of approximating the density function of states using other known PDFs is referred to as kernel density estimation (KDE) [52]. For example, let us consider the set of PDF described of a r-parameter set, given by, $\mathbf{P}(\alpha_1, \alpha_2, \ldots, \alpha_r)$, where $\{\alpha_i\}_{i=1}^{r}$ are the set of parameters. Using GMM this PDF can be represented as,

$$\mathbf{P}(\alpha_1, \alpha_2, \ldots, \alpha_r) = \sum_{j=1}^{\infty} \beta_j \mathcal{N}(\mu_j, \Sigma_j) \tag{2.5}$$

where β_j are constants and $\mathcal{N}(\mu_j, \Sigma_j)$ are Gaussian PDFs with μ_j and Σ_j being mean and covariance, form a basis for representing such PDFs.

For estimation purposes, if the initial density is given by $\mathbf{P}(\alpha_1, \alpha_2, \ldots, \alpha_r)$, we start be representing it in GMM framework using Eqn. (2.5). The expansion in Eqn. (2.5) is truncated to N_t terms. Each Gaussian basis PDF is propagated using Eqn. (2.3a) & Eqn. (2.3b) to get the prior mean and covariance, and the update step involves using Eqn. (2.4a) through Eqn. (2.4c) for each Gaussian PDF. The posterior state PDF is obtained by using a GMM model with posterior mean and covariance from update step parameterizing each Gaussian PDF. The essential steps for state estimation with non-Gaussian initial PDF is given in Table I. Detailed discussion of estimation using GMM can be found in ref. [82].

Table I.: Algorithm for linear non-Gaussian filtering

Step	**Equations**
Initialization	$\mathbf{P}(\alpha_1, \alpha_2, \ldots, \alpha_r)(t_0) = \sum_{j=1}^{N_t} \beta_j \mathcal{N}(\mu_j(t_0), \Sigma_j(t_0))$.
Propagation	Propagate $\mu_j(t_{k\|k})$ and $\Sigma_j(t_{k\|k})$ using Eqn. (2.3a) & Eqn. (2.3b) to get $\mu_j(t_{k+1\|k})$ and $\Sigma_j(t_{k+1\|k})$.
Update	Update $\mu_j(t_{k+1\|k})$ and $\Sigma_j(t_{k+1\|k})$ using Eqn. (2.4a) through Eqn. (2.4c) to get $\mu_j(t_{k+1\|k+1})$ and $\Sigma_j(t_{k+1\|k+1})$.
Final PDF	$\mathbf{P}(\alpha_1, \alpha_2, \ldots, \alpha_r)(t_{k+1}) = \sum_{j=1}^{N_t} \beta_j \mathcal{N}(\mu_j(t_{k+1\|k+1}), \Sigma_j(t_{k+1\|k+1}))$.

B. Sequential State Estimation for Nonlinear Systems

In this section we will introduce some popular methods of state estimation for nonlinear systems. For nonlinear systems, there exist no estimator that is optimal with respect to the established optimality criteria [73]. Hence, all the algorithms described henceforth yield suboptimal solution, using some approximation methods. To judge the performance of these solutions, there are several metrics that have become useful. Towards the end of this chapter we will discuss in brief some of the criteria that are used as a metric to judge the performance of the estimation algorithm, through an example. We would begin with Kalman filter based methods and then go on to describe more robust methods which are also known as sequential Monte Carlo (SMC) methods.

To give a generic flavor to the problem we define the dynamical system and the observation model in a way such that they can be used in subsequent sections. Let us consider two real valued functions which are at least once continuously differentiable, $f : (x, u) \to \mathbb{R}^n$ and $h : (x, u) \to \mathbb{R}^m$ where $f, h \in \mathcal{C}_1$, where x, u are states and controls

and are given by, $x : t \to \mathbb{R}^n$ and $u : t \to \mathbb{R}^p$, respectively. Let the measurements be given by $\tilde{y} : t \to \mathbb{R}^m$. Let us also consider a class of zero mean δ-correlated Gaussian noise $w(t)$ and $v(t)$ which are the process and the measurement noise respectively, and have the autocovariance $Q\delta(t - t')$ and $R\delta(t - t')$, respectively; $t, t' \in \mathbb{R}^+$. For sake of simplicity, the Q and R matrices are assumed to be constant. The dynamical system and the measurement equations are given by,

$$\dot{x}(t) = f(x, u) + w(t) \tag{2.6a}$$

$$\tilde{y}(t) = h(x, u) + v(t) \tag{2.6b}$$

As before, we consider the continuous-discrete formulation of the filter equations, with measurements arriving at discrete intervals of time $t_1, t_2, \ldots, t_k, t_{k+1}, \ldots$. In such case Eqn. (2.6a) & Eqn. (2.6b) are given by,

$$\dot{x}(t) = f(x, u) + w(t) \tag{2.7a}$$

$$\tilde{y}_k = h(x_k, u_k) + v_k \tag{2.7b}$$

Equation (2.7a) & Eqn. (2.7b) will be used throughout the development of nonlinear filtering methods.

1. Extended Kalman Filter

In estimation theory, the extended Kalman filter (EKF) is the nonlinear version of the Kalman filter where the stochastic dynamics is linearized about the current estimate. The linearized dynamics is propagated forward in time to get the sequence of evolving mean and covariance.

EKF assumes linear behavior of the nonlinear system for sufficiently small propa-

gation time, and hence uses the linearized dynamics to propagate uncertainty. Also it assumes Gaussian PDF evolution, hence mean and covariance of the states are propagated, and updated using Kalman update law. Here, as in the case of Kalman filter, the mean is the state estimate. The mean and covariance propagation equations are given by,

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) \text{ with, } \hat{x}(t_k) = \hat{x}_{k|k} \tag{2.8a}$$

$$\dot{P}(t) = F(t)P(t) + P(t)F(t)^\top + Q \text{ with, } P(t_k) = P_{k|k} \tag{2.8b}$$

which are propagated from $t \in [t_k, t_{k+1}]$.

The update equations are same as that for the Kalman filter which are given by Eqn. (2.4a) through Eqn. (2.4c), where $F(t)$ and $H(t)$ are Jacobians given by,

$$F(t) = \frac{\partial f}{\partial x}|_{\hat{x}(t),u(t)} \;\; H(t) = \frac{\partial h}{\partial x}|_{\hat{x}(t),u(t)}$$

f and h are functions that were defined in Eqn. (2.6a) & Eqn. (2.6b).

Due to linearization, EKF has been found to accrue errors if the propagation times are long [83]. Hence for systems where the measurement update are infrequent, estimation errors are observed to be divergent. Moreover, like its linear counterpart EKF assumes the PDF evolution is Gaussian which is not always true for a nonlinear system. It has also been observed that if the initial error estimates are large, the covariance matrix underestimates the true covariance and the results of EKF are found to be inconsistent. Though this error can be corrected marginally by selecting R and Q properly.

2. Unscented Kalman Filter

The unscented Kalman filter (UKF) uses a deterministic sampling technique known as the unscented transform to pick a minimal set of sample points, called sigma points,

around the mean [7]. These sigma points are then propagated through the nonlinear functions, from which the mean and covariance of the estimate are then recovered. It has been shown that the result is a filter which more accurately captures the true mean and covariance, than EKF [84]. In addition, this technique removes the requirement to explicitly calculate the Jacobians, which, for complex functions, can be a difficult task in itself.

We will assume the same nonlinear estimation setting as given in Eqn. (2.6a) & Eqn. (2.6b). In the prediction step of UKF, the estimated state and covariance are augmented with the mean and covariance of the process noise, i.e.

$$x^a_{k|k} = [\hat{x}^T_{k|k} \quad E[w^T_{k+1}]\ ]^T \tag{2.9a}$$

$$P^a_{k|k} = \begin{bmatrix} P_{k|k} & 0 \\ 0 & Q \end{bmatrix} \tag{2.9b}$$

A set of $2L+1$ sigma points is derived from the augmented state and covariance where L is the dimension of the augmented state. The sigma points are given by,

$$\chi^1_{k|k} = x^a_{k|k} \tag{2.10a}$$

$$\chi^i_{k|k} = x^a_{k|k} + \left(\sqrt{(L+\lambda)P^a_{k|k}}\right)_i, \ \ i = 2 \dots L+1 \tag{2.10b}$$

$$\chi^i_{k|k} = x^a_{k|k} - \left(\sqrt{(L+\lambda)P^a_{k|k}}\right)_{i-L}, \ \ i = L+2, \dots, 2L+1 \tag{2.10c}$$

where $\left(\sqrt{(L+\lambda)P^a_{k|k}}\right)_i$ is the i^{th} column of the matrix square root of $(L+\lambda)\,P^a_{k|k}$. The quantity λ, is defined as,

$$\lambda = \alpha^2\,(L+\kappa) - L \tag{2.11}$$

where α and κ control the spread of the sigma points. Normal value of $\alpha = 10^{-3}$ and $\kappa = 1$. However, one may change these constants according to application.

The sigma points are propagated using the same equation as given in Eqn. (2.6a), from time $[t_k, t_{k+1}]$ to get $\chi^i_{k+1|k}$. Let $g(x, u, w) = f(x, u) + w(t)$, then

$$\dot{\chi}(t)^i = g(\chi(t)^i) \quad i = 1 \dots 2L + 1, \quad \text{with, } \chi(t_k)^i = \chi^i_{k|k} \tag{2.12}$$

The weighted sigma points are recombined to produce the predicted prior state and covariance.

$$\hat{x}_{k+1|k} = \sum_{i=1}^{2L+1} W^i_s \chi^i_{k+1|k} \tag{2.13a}$$

$$P_{k+1|k} = \sum_{i=1}^{2L+1} W^i_c \, [\chi^i_{k+1|k} - \hat{x}_{k+1|k}][\chi^i_{k+1|k} - \hat{x}_{k+1|k}]^T \tag{2.13b}$$

where the weights for the state and covariance are given by,

$$W^1_s = \frac{\lambda}{L + \lambda} \tag{2.14a}$$

$$W^1_c = \frac{\lambda}{L + \lambda} + (1 - \alpha^2 + \beta) \tag{2.14b}$$

$$W^i_s = W^i_c = \frac{1}{2(L + \lambda)} \tag{2.14c}$$

β is a constant related to the distribution of the states. For example, if the underlying distribution is Gaussian then $\beta = 2$ is optimal.

The predicted prior state and covariance are augmented as previously, except now the estimated state vector and the covariance matrix are augmented by the mean and covariance of the measurement noise.

$$x^a_{k+1|k} = [\hat{x}^T_{k+1|k} \quad E[v^T_{k+1}] \ ]^T \tag{2.15a}$$

$$P^a_{k+1|k} = \begin{bmatrix} P_{k+1|k} & 0 \\ 0 & R \end{bmatrix} \tag{2.15b}$$

As in the case of propagation step, a set of $2L+1$ sigma points is derived from the augmented state and covariance (in Eqn. (2.15a) & Eqn. (2.15b)) where L is the dimension of the augmented state.

$$\chi^1_{k+1|k} = x^a_{k+1|k} \tag{2.16a}$$

$$\chi^i_{k+1|k} = x^a_{k+1|k} + \left(\sqrt{(L+\lambda)P^a_{k+1|k}}\right)_i, \quad i = 2 \dots L+1 \tag{2.16b}$$

$$\chi^i_{k+1|k} = x^a_{k+1|k} - \left(\sqrt{(L+\lambda)P^a_{k+1|k}}\right)_{i-L}, \quad i = L+2, \dots, 2L+1 \tag{2.16c}$$

One can also use the sigma points received after propagation in Eqn. (2.12), given by,

$$\chi_{k+1|k} := [\chi^T_{k+1|k} \quad E[v^T_{k+1}]\ ]^T \pm \sqrt{(L+\lambda)R^a} \tag{2.17}$$

where,

$$R^a = \begin{bmatrix} 0 & 0 \\ 0 & R \end{bmatrix}$$

Let, $\bar{h}(\chi_k) = h(x_k, u_k) + v_k$ be the discrete observation process. Then we get the weighted sigma points using the following equation,

$$\gamma^i_{k+1} = \bar{h}(\chi^i_{k+1|k}) \quad i = 1 \dots 2L+1 \tag{2.18}$$

The predicted measurements, and their covariance, and also the state covariance are then received from the weighted sigma points using the equation,

$$\hat{y}_{k+1} = \sum_{i=1}^{2L+1} W^i_s \gamma^i_{k+1} \tag{2.19a}$$

$$P_{y_{k+1}y_{k+1}} = \sum_{i=1}^{2L+1} W^i_c\ [\gamma^i_{k+1} - \hat{y}_{k+!}][\gamma^i_{k+1} - \hat{y}_{k+1}]^T \tag{2.19b}$$

$$P_{x_{k+1}y_{k+1}} = \sum_{i=1}^{2L+1} W^i_c\ [\chi^i_{k+1|k} - \hat{x}_{k+1|k}][\gamma^i_{k+1} - \hat{y}_{k+1}]^T \tag{2.19c}$$

The Kalman gain K_k is then computed using the equation,

$$K_{k+1} = P_{x_{k+1}y_{k+1}} P^{-1}_{y_{k+1}y_{k+1}} \tag{2.20}$$

The posterior state and covariance are then received using the Kalman gain and the measurements.

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}(y_{k+1} - \hat{y}_{k+1}) \tag{2.21a}$$

$$P_{k+1|k+1} = P_{k+1|k} - K_{k+1} P_{y_{k+1}y_{k+1}} K^T_{k+1} \tag{2.21b}$$

Though UKF provides us with a method, where we can avoid the disadvantages due to linearization of dynamics but its main drawback is assumption of Gaussian PDF evolution (i.e. use of Kalman update law). In most real-world situations PDF evolution is non-Gaussian and UKF has been observed to perform unsatisfactorily in such cases. The state estimated covariance and the true covariance don't match and so the estimator becomes inconsistent. Hence, a state estimation methodology where the PDF evolution is assumed to be non-Gaussian, is desired.

3. Particle Filters

Particle filters (PF), also known as sequential Monte Carlo methods (SMC), are sophisticated state estimation techniques based on Monte Carlo simulations. They are based upon *importance sampling theorem* [11], where we draw random samples from a "proposal distribution" based on their "weights", and propagate using Eqn. (2.6a) [10]. Particle filters are often used in scenarios where the EKF or UKF fail, with the advantage that, with sufficiently large number of particles, they approach the Bayesian optimal estimate, so they can be made more accurate than either the EKF or UKF. However, when the simulated sample is not sufficiently large, they

might suffer from sample impoverishment. The approaches can also be combined by using a version of the Kalman filter as a proposal distribution for the particle filter.

Let us consider the system given in Eqn. (2.6a) & Eqn. (2.6b). Let the initial states have the PDF $P(x(t=0))$. We follow the following steps in particle filtering.

a. Step 1: Initialization of the Filter

We draw N particles from the domain of initial state $x(t=0)$ with replacement, which are given by $x_{0,i},\ i=1,2,\cdots,N$, where $p(x(t_0)=x_{0,i})$ is the probability of selection of the i^{th} particle. The initial weights are given by

$$w_{0,i} = \frac{p(x(t_0)=x_{0,i})}{\sum_{j=1}^{N} p(x(t=0)=x_{0,j})} \tag{2.22}$$

The state estimate at time t_0 is the weighted mean of all particles, i.e. $\hat{x}_0 = \sum_{i=1}^{N} w_{0,i}x_{0,i}$. We now perform steps 2 to 4 recursively starting from $k=1$.

b. Step 2: Propagation

We now get $x_{k|k-1,i}$ for each particle i, by integrating Eqn. (2.6a) over the interval $[t_{k-1}, t_k]$, with initial states as $x_{k-1|k-1,i}$.

The particles $x_{k|k-1,i}$, represent weighted sample which is received from the prior PDF $p(x(t_k)|x(t_{k-1}))$. Generally it is very difficult to sample from the prior PDF as its exact analytical representation is unknown. Let us assume a "proposal" PDF $\pi(x(t_k)|x(t_{k-1}))$, which is close to the prior PDF and easy to sample from. We sample N particles from the proposal PDF, which we represent as $x_{k|k,i}$

c. Step 3: Update

We update the weights $w_{k-1,i}$ using Bayesian update rule [78]. We first construct the likelihood function, $p_i(\tilde{y}_k|x(t_k) = x_{k|k,i})$ for each particle i, using the Gaussian measurement noise, and the sensor model as shown in Eqn. (2.6b). It is defined as

$$p_i(\tilde{y}_k|x(t_k) = x_{k|k,i}) = \frac{1}{\sqrt{(2\pi)^m|R|}} e^{-\frac{1}{2}(\tilde{y}_k - h(x_{k|k,i}))^T R^{-1}(\tilde{y}_k - h(x_{k|k,i}))} \tag{2.23}$$

where $|R|$ is the determinant of measurement noise covariance matrix.

The weights are then updated up to a normalizing constant using the equation,

$$\hat{w}_{k,i} = \frac{p_i(\tilde{y}_k|x(t_k) = x_{k|k,i})p(x(t_k)|x(t_{k-1}))}{\pi(x(t_k)|x(t_{k-1}))} w_{k-1,i} \tag{2.24}$$

Note if the proposal density is the prior then Eqn. (2.24) reduces to

$$\hat{w}_{k,i} = p_i(\tilde{y}_k|x(t_k) = x_{k|k,i})w_{k-1,i}$$

The weights are then normalized to get the final weights

$$w_{k,i} = \frac{\hat{w}_{k,i}}{\sum_{i=1}^{N} \hat{w}_{k,i}} \tag{2.25}$$

The above method of using a proposal density to obtain the unbiased sample is often called importance sampling [11].

d. Step 4: State Estimate

We then statistically approximate the state estimate as, (e.g. [15, 16])

$$\hat{x}_k = \sum_{i=1}^{N} w_{k,i} x_{k|k,i} \tag{2.26}$$

It should be noted that in the limit of infinitely large N, particle filter gives us

asymptotically exact estimate of the state.

e. Resampling

In most practical applications, a large number of the weights, $w_{k,i}$ become negligible after certain number of recursive steps. This phenomenon is called degeneracy. Hence a large computational effort is wasted in updating weights making little contribution towards state estimate. A measure of degeneracy at step k is the effective sample size [13, 15], given by

$$N_e = \frac{1}{\sum_{i=1}^{N} w_{k,i}^2}$$

If all but one weight is zero then $N_e = 1$ indicating degeneracy. We set a threshold value, N_t for effective number of particles, and resample whenever $N_e < N_t$. The resampling is done in following manner.

1. Draw N particles from the current particle set, $x_{k|k,i}$ with probability of selection as $w_{k,i}$. Replace the current particle set with the new one.

2. Set $w_{k,i} = 1/N$ for $i = 1, 2, \cdots, N$.

Although resampling step eliminates degeneracy, it can artificially reduce the estimated state variance thus giving erroneous state estimate.

Note that the algorithm presented above is one of the many particle filtering algorithms in use, but is the most common one. There are several variants of the particle filtering algorithm depending on application. In next section, we will introduce in brief, some of the particle filtering algorithms that are most commonly used.

4. Other Sequential Monte Carlo Methods

Researchers have developed several SMC methods that would suit their particular application. These algorithms use Monte Carlo methods for propagation and then use Bayesian update. They vary within themselves by the choice of proposal PDF or the resampling method used. In this section we will discuss in brief some of the SMC methods that are commonly used by practitioners. We will just introduce them and explain the difference from the algorithm presented in the previous section, without going into the details of each algorithm.

The algorithm presented in the previous section is often termed as *generic particle filter*. There are several variants to this filtering method, namely in the resampling step. All of them are known by the name of generic PF. For example Arulampalam *et al.* , presents a new algorithm, which uses a MCMC based resampling technique [9]. If we eliminate the resampling step entirely, the algorithm is called *sequential importance sampling* (SIS) [85], and if we resample at each step, disregarding the effective number of particles, the resulting filter is known as *sequential importance resampling* (SIR) or bootstrap filter [8].

Resampling is seen to be a major step in particle filtering algorithms as most particle filters developed suffer from sample impoverishment [15]. Hence, a major effort has been put to make the resampling step more robust to increase diversity amongst particles. Hence particle filters like *regularized particle filter* and *MCMC Move step* particle filters have been developed, which use better resampling techniques. But particle filters, being a simulation based method suffer from the curse of dimensionality, i.e. the computational cost increases exponentially with increase in state dimension [22]. To solve this problem, particle filters based on Rao-Blackwellization have been developed to partially solve the estimation problem analytically [23]. A description

of the SMC techniques commonly in use can be found in ref. [15].

C. A Simple Example

We consider a Duffing oscillator system with dynamics and the discrete measurements given by the equations,

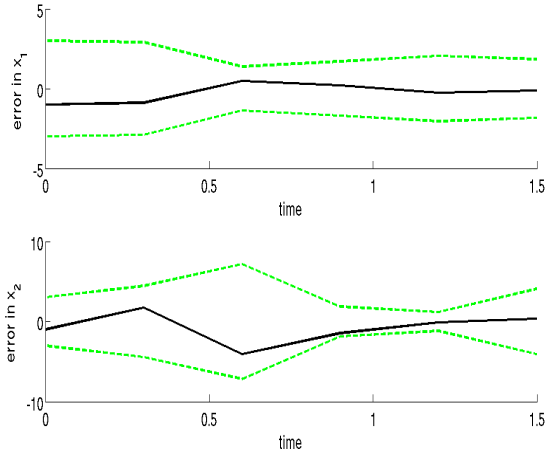
$$\ddot{x}(t) = -x(t) - \frac{1}{4}x(t)^3 - \dot{x}(t) + w(t) \tag{2.27a}$$

$$\tilde{y}_k = x_k^2 + \dot{x}_k^2 + v_k \tag{2.27b}$$

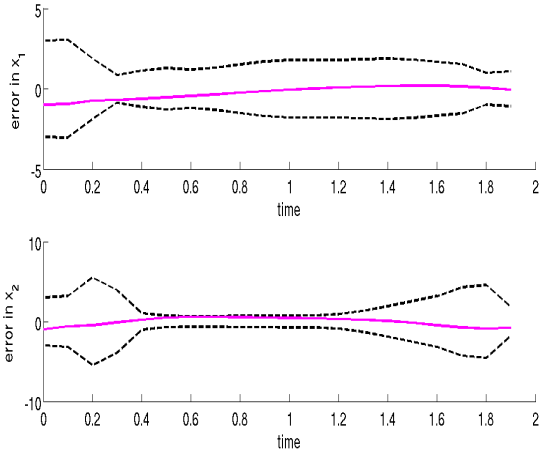
where $w(t)$ and $v(t)$ are zero mean process noise with autocorrelation $Q = 6 \times 10^{-2}$ and $R = 6 \times 10^{-1}$, respectively. The initial states of the system are assumed to have Gaussian PDF with mean and covariance as $[1, 1]$ and $\text{diag}(1, 1)$, respectively.

Figure 1 shows plots for estimation error and $\pm 3\sigma$ limits for EKF, UKF and particle filter (PF), when used to estimate states of the system in Eqn. (2.27a) & Eqn. (2.27b). The solid line represents error in estimation, i.e. how close are the estimated states to the actual states of the system. The dashed lines represent $\pm 3\sigma$ confidence intervals or $\pm 3\sigma$ limits. This refers to the confidence level of the estimated states. It says that the estimation algorithm is 6σ percent confident that the estimation error will lie within the limits. For a Gaussian distribution, this is a very high number which is equal to 97.3% [73] . Hence, normal intuition suggests that the estimation error should lie within the $\pm 3\sigma$ limits if the PDF propagation is assumed to be Gaussian. However, if Gaussian propagation is not assumed, the plots can be inconclusive in some cases.

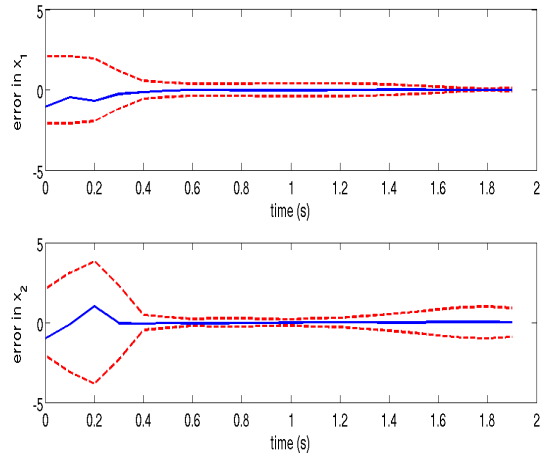
In Fig. 1, we assume that the update interval of each measurement is 0.1s. We observe that the errors are within $\pm 3\sigma$ limits, hence we can conclude that all the three filters are successful in prediction of uncertainty. Moreover we can see that the $\pm 3\sigma$



(a) EKF.



(b) UKF.



(c) Particle filter.

Fig. 1.: Plot of estimation errors and $\pm 3\sigma$ limits for a) EKF b) UKF c) particle filter for the system in Eqn. (2.27a) & Eqn. (2.27b). The solid lines are estimation errors and the dashed lines are $\pm 3\sigma$ limits.

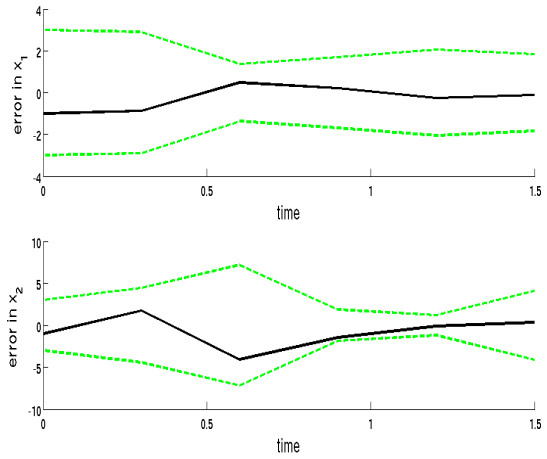
limits and estimation error of the PF converge faster than EKF and UKF. Hence for the given system PF is superior than EKF or UKF. This is because of the dynamics is nonlinear and the evolution of PDF is non-Gaussian.

In Fig. 2, we increase the update interval to 0.3s. We can see that the performance of PF is better than EKF and UKF. The poor performance of EKF and UKF is more conspicuous in this case as the linear Gaussian assumption doesn't hold for sufficiently long propagation times. Hence for a nonlinear system where PDF evolution is Gaussian particle filters perform the best. This is in agreement with the theory about sequential state estimation.

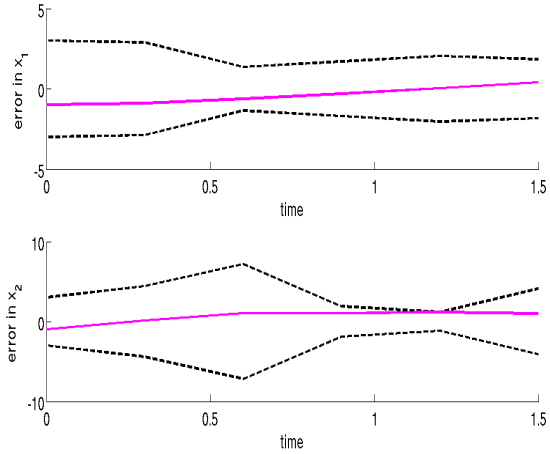
As mentioned earlier, $\pm 3\sigma$ limits do not provide conclusive evidence of efficacy of an estimation algorithm if the PDF evolution is non-Gaussian. In such cases Cramer-Rao bounds can be used as a metric to judge their effectiveness [9]. Cramer-Rao lower bound (CRLB) gives us a bound for covariance minimization of an estimation algorithm. It says that, given a suboptimal estimation algorithm, minimizing variance, the posterior variance of that algorithm is lower bounded by CRLB. So better the estimation algorithm closer the minimum variance solution is to CRLB. In Fig. 3 we plot the square root of the difference between the Cramer-Rao lower bounds and the variance for each estimator compared in this section for the system in Eqn. (2.27a) & Eqn. (2.27b). Clearly it can be seen that PF has a smaller difference than the other estimators assuming Gaussian PDF evolution. This shows the suitability of PF in a nonlinear non-Gaussian estimation setting.

D. Summary of the Chapter

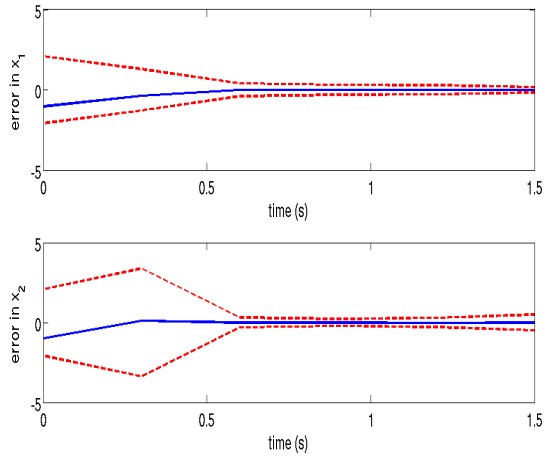
In this chapter, we have introduced sequential state estimation techniques for linear and nonlinear systems, assuming both Gaussian and non-Gaussian PDF evolution.



(a) EKF.



(b) UKF.



(c) Particle filter.

Fig. 2.: Plot of estimation errors and $\pm 3\sigma$ limits for a) EKF b) UKF c) particle filter for the system in Eqn. (2.27a) & Eqn. (2.27b). The measurement update interval is 0.3s. The solid lines are estimation errors and the dashed lines are $\pm 3\sigma$ limits.

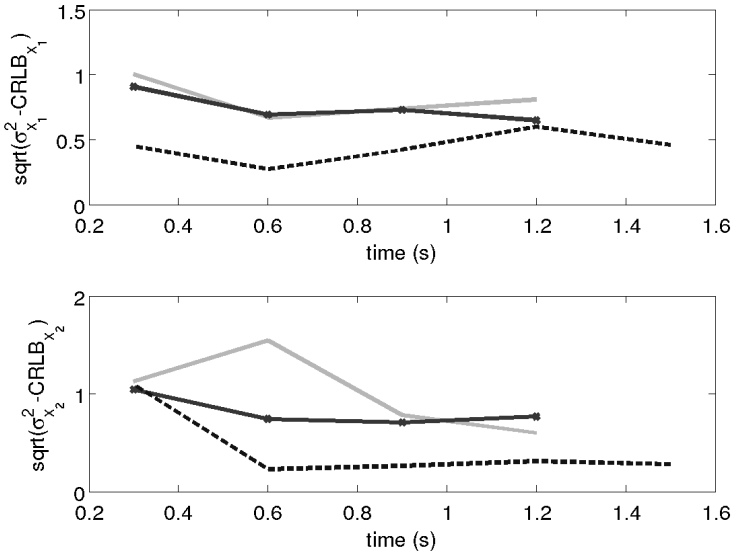


Fig. 3.: Plot of square root of difference between evolving variance of each state and CRLB. The solid line is EKF the starred line is UKF and the dashed line is particle filter.

We have shown through an example that nonlinear non-Gaussian estimation techniques are better for systems which are nonlinear in nature. In the following chapters, we will discuss about the estimation algorithms proposed in this dissertation. For the sake of comparison, we will continuously refer to the algorithms described in this chapter, and show how the proposed algorithms perform when compared to the estimation techniques described in this chapter.

CHAPTER III

POLYNOMIAL CHAOS*

Polynomial chaos (PC) is a parametric method based on using orthogonal functionals to represent random processes that are solutions of dynamic systems with uncertainties. It utilizes families of orthogonal polynomials, which we will refer to as polynomial chaoses, to approximate the both the functions of random variables which appear in the equations of motion for a dynamic system as well as the actual solution. In this chapter, we define the structure of these orthogonal polynomials and present some of their properties, which will be applied to estimate states of dynamical systems having uncertainty.

A. Generalized Polynomial Chaos Theory

Let $(\Omega, \mathcal{F}, \mathcal{M})$ be a probability space, where Ω is the sample space, $\mathcal{F}$ is the σ-algebra of the subsets of Ω, and $\mathcal{M}$ is the probability measure. Let $\Delta(\omega) = (\Delta_1(\omega), \cdots, \Delta_d(\omega)) : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}^d, \mathcal{B}^d)$ be an $\mathbb{R}^d$-valued continuous random variable, where $d \in \mathbb{N}$, and $\mathcal{B}^d$ is the σ-algebra of Borel subsets of $\mathbb{R}^d$. A general second order process $X(\omega) \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathcal{M})$ can be expressed in polynomial chaos framework as

$$X(\omega) = \sum_{i=0}^{\infty} x_i \phi_i(\Delta(\omega)), \tag{3.1}$$

where ω is the random event and $\phi_i(\Delta(\omega))$ denotes the generalized polynomial chaos (gPC) basis function of degree i, in terms of the random variables $\Delta(\omega)$. Henceforth, Δ will be use to represent $\Delta(\omega)$.

---

*Reprinted from "Nonlinear Estimation of Hypersonic State Trajectories in Bayesian Framework with Polynomial Chaos" by P. Dutta, R. Bhattacharya, 2010. *AIAA Journal of Guidance Control and Dynamics*, vol. 33, no. 6, pp. 1765–1778, Copyright [2010] by P. Dutta & R. Bhattacharya.

1. Wiener-Askey Orthogonal Polynomials

To approximate a stochastic process, a set of orthogonal polynomials will be employed. In this section, we will present an overview of how to generate such polynomials for the gPC framework. Given the random variable Δ with probability density function (PDF), $p(\Delta)$, let $v = [1, \Delta, \Delta^2, \ldots, \infty]^T$. The family of orthogonal basis functions $\{\phi_i(\Delta)\}$ are given by,

$$\phi_0(\Delta) = v_0 \tag{3.2a}$$

$$\phi_i(\Delta) = v_i - \sum_{k=0}^{i-1} \frac{\langle v_i, \phi_k(\Delta)\rangle}{\langle \phi_k(\Delta), \phi_k(\Delta)\rangle}\phi_k(\Delta) \quad i = 1, \ldots, \infty, \tag{3.2b}$$

where

$$\langle \phi_i, \phi_j \rangle = \int_{\mathbf{D}_\Delta} \phi_i \phi_j p(\Delta)\, d\Delta, \tag{3.3}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product with respect to the weight function $p(\Delta)$, and $\mathbf{D}_\Delta$ is the domain of the random variable Δ. Note that the weight function for the inner product here is same as the PDF of Δ.

For example, we take a scalar case i.e. $d = 1$, let the PDF of Δ be a standard normal PDF, then

$$p(\Delta) = \frac{1}{(2\pi)} e^{-\frac{1}{2}\Delta^T\Delta},$$

where $\Delta \in \mathbb{R}$. To find the orthogonal polynomials $\{\phi_i(\Delta)\}_{i=1}^{\infty}$, we need to employ the

scheme given in Eqn. (3.2a) & Eqn. (3.2b). Hence,

$$\begin{aligned}
\phi_0(\Delta) &= 1 \\
\phi_1(\Delta) &= \Delta - \frac{\langle \Delta, 1 \rangle}{\langle 1, 1 \rangle} \times 1 = \Delta \\
\phi_2(\Delta) &= \Delta^2 - \frac{\langle \Delta^2, 1 \rangle}{\langle 1, 1 \rangle} \times 1 - \frac{\langle \Delta^2, \Delta \rangle}{\langle \Delta, \Delta \rangle} \times \Delta = \Delta^2 - 1 \\
&\vdots \\
\phi_i(\Delta) &= \Delta^i - \sum_{k=0}^{i-1} \frac{\langle \Delta^i, \phi_k(\Delta) \rangle}{\langle \phi_k(\Delta), \phi_k(\Delta) \rangle} \times \phi_k(\Delta).
\end{aligned}$$

The procedure in Eqn. (3.2a) & Eqn. (3.2b) is the classical *Gram-Schmidt Orthogonalization* [86]. The orthogonal polynomials thus obtained are the members of the Askey-scheme of polynomials [30], which form a complete basis in the Hilbert space determined by their corresponding support. Table II summarizes the correspondence between the orthogonal polynomials for a given PDF of Δ [29].

Table II.: Correspondence between choice of polynomials and given distribution of Δ (Xiu and Karniadakis, 2002)

PDF of Δ	**$\phi_i(\Delta)$ of the Wiener-Askey Scheme**
Gaussian	Hermite
Uniform	Legendre
Gamma	Laguerre
Beta	Jacobi

B. Approximation of the Solution of Ordinary Differential Equations with Uncertainty

A dynamical system of the form $\dot{x} = f(x, \Delta)$, where $x \in \mathbb{R}^n$ and random variable $\Delta \in \mathbb{R}^d$, representing uncertainty in initial states and parameters, can be solved in the gPC framework in the following manner. Assume solution of the differential equation to be $x(t, \Delta)$. For second order processes, the solution for every component of $x \in \mathbb{R}^n$ can be approximated as

$$\hat{x}_i(t, \Delta) = \sum_{j=0}^{N} x_{ij}(t)\phi_j(\Delta);\ i = 1, \cdots, n. \tag{3.4}$$

The above series is truncated after $N+1$ terms, which is determined by the dimension d of Δ and the order r of the orthogonal polynomials $\{\phi_j\}$, satisfying $N + 1 = (d+r)!/d!r!$. This expression gives the number of terms in a sequence of multi-variate polynomials up to order r with d variables.

Substituting the approximate solution into equation of the dynamical system results in errors which are given by,

$$e_i = \dot{\hat{x}}_i - f_i(\hat{x}, \Delta);\ i = 1, \cdots, n.$$

The approximation in Eqn. (3.4) is optimal in the $\mathcal{L}_2$ sense when the projection of the errors on the orthogonal basis functions are zero, i.e.,

$$\langle e_i(t, \Delta), \phi_j(\Delta) \rangle = 0, \tag{3.5}$$

for $j = 0, \cdots, N;\ i = 1, \cdots, n$. Equation (3.5) results in the following $n(N + 1)$ *deterministic* ordinary differential equations

$$\dot{x}_{ik} = \frac{\int_{\mathbf{D}_\Delta} f\left(\sum_{j=0}^{N} x_{ij}(t)\phi_j(\Delta), \Delta\right) \phi_k(\Delta) p(\Delta) d\Delta}{\int_{\mathbf{D}_\Delta} \phi_k^2(\Delta) p(\Delta) d\Delta}, \tag{3.6}$$

for $i = 1, \cdots, n$ and $k = 0, \cdots, N$. Therefore, the uncertain dynamics in $\mathbb{R}^n$ has been transformed into deterministic dynamics in $\mathbb{R}^{n(N+1)}$. Let us represent $X_{pc} = [x_{10} \cdots x_{1N}\, x_{20} \cdots x_{2N} \cdots x_{n0} \cdots x_{nN}]^T$. Then Eqn. (3.6) can be written in a compact form as

$$\dot{X}_{pc} = F_{pc}(X_{pc}), \tag{3.7}$$

where $F_{pc}(X_{pc})$ represents the right hand side of Eqn. (3.6). Equation (3.7) can be solved using algorithms for ordinary differential equation, to obtain the approximate stochastic response of the system under consideration.

This method of obtaining Eqn. (3.7) through Eqn. (3.6) is referred to as intrusive method, and is difficult to compute when $f(\hat{x}, \Delta)$ is a non-polynomial function [87]. For such cases, it is better to apply gPC on a suitable polynomial approximation of $f(\hat{x}, \Delta)$.

1. Getting the Moments of the States

Given the gPC expansion we can get the moments of the states of the dynamical system using the gPC coefficients. For example, using gPC, the mean of any state is

given by,

$$\begin{aligned}
\mathbb{E}\left[x(t)\right] &= \int_{\mathbf{D}_\Delta} x(t,\Delta)p(\Delta)d\Delta \\
\Rightarrow \mathbb{E}\left[x(t)\right] &= \int_{\mathbf{D}_\Delta} \sum_{i=1}^{N} x_i(t)\phi_i(\Delta)p(\Delta)d\Delta \\
&= \sum_{i=1}^{N} x_i(t) \int_{\mathbf{D}_\Delta} \phi_i(\Delta)p(\Delta)d\Delta \\
\Rightarrow \mathbb{E}\left[x(t)\right] &= [x_1, x_2, \ldots, x_N] \begin{bmatrix} \langle\phi_1(\Delta)\rangle \\ \langle\phi_2(\Delta)\rangle \\ \vdots \\ \langle\phi_N(\Delta)\rangle \end{bmatrix}.
\end{aligned} \tag{3.8}$$

We use a similar methodology to get the second moment of the states.

$$\begin{aligned}
\mathbb{E}\left[x(t)^2\right] &= \int_{\mathbf{D}_\Delta} x^2(t,\Delta)p(\Delta)d\Delta \\
\Rightarrow \mathbb{E}\left[x(t)^2\right] &= \int_{\mathbf{D}_\Delta} \sum_{i=1}^{N}\sum_{j=1}^{N} x_i(t)x_j(t)\phi_i(\Delta)\phi_j(\Delta)p(\Delta)d\Delta \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} x_i(t)x_j(t) \int_{\mathbf{D}_\Delta} \phi_i(\Delta)\phi_j(\Delta)p(\Delta)d\Delta \\
\Rightarrow \mathbb{E}\left[x(t)^2\right] &= [x_1(t), \ldots x_N(t)] \begin{bmatrix} \langle\phi_1(\Delta)\phi_1(\Delta)\rangle & \ldots & \langle\phi_1(\Delta)\phi_N(\Delta)\rangle \\ \langle\phi_2(\Delta)\phi_1(\Delta)\rangle & \ldots & \langle\phi_2(\Delta)\phi_N(\Delta)\rangle \\ \vdots & \ddots & \vdots \\ \langle\phi_N(\Delta)\phi_1(\Delta)\rangle & \ldots & \langle\phi_N(\Delta)\phi_N(\Delta)\rangle \end{bmatrix} \begin{bmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{bmatrix}.
\end{aligned} \tag{3.9}$$

The third and higher order moments of the states can be found similarly using the gPC coefficients, and the inner product of basis functions. Comprehensive derivation of moments with explanation of each step can be found in [88].

2. A Simple Example

We will now show, using an example, how gPC is used to solve ordinary differential equations with uncertainty. Let us consider a Duffing oscillator with the following dynamics,

$$\dot{x}_1 = x_2 \tag{3.10}$$

$$\dot{x}_2 = -x_1 - \frac{1}{4}x_2 - x^3. \tag{3.11}$$

Let the above dynamical system have initial state uncertainty, $\Delta := [x_1(t=0), x_2(t=0)]$ with PDF $p(\Delta)$. Let the orthogonal polynomials found using Eqn. (3.2a) & Eqn. (3.2b) be $\phi_i(\Delta)$. Then the states at any time t can be represented in gPC framework as

$$x_1(t) = \sum_{i=1}^{N} x_{1i}(t)\phi_i(\Delta) \tag{3.12a}$$

$$x_2(t) = \sum_{i=1}^{N} x_{2i}(t)\phi_i(\Delta). \tag{3.12b}$$

Substituting the expressions in Eqn. (3.12a) & Eqn. (3.12b) in Eqn. (3.10) & Eqn. (3.11), we get

$$\sum_{i=1}^{N} \dot{x}_{1i}(t)\phi_i(\Delta) = \sum_{i=1}^{N} x_{2i}(t)\phi_i(\Delta) \tag{3.13a}$$

$$\sum_{i=1}^{N} \dot{x}_{1i}(t)\phi_i(\Delta) = -\sum_{i=1}^{N} x_{1i}(t)\phi_i(\Delta) - \frac{1}{4}\sum_{i=1}^{N} x_{2i}(t)\phi_i(\Delta) - \sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N} x_{1j}(t)x_{1k}(t)x_{1l}(t)\phi_j(\Delta)\phi_k(\Delta)\phi_l(\Delta). \tag{3.13b}$$

Taking inner product with respect to $\phi_i(\Delta)$ in both sides of Eqn. (3.13a) and Eqn. (3.13b) and using orthogonality of $\phi_i(\Delta)$ we have

$$\dot{x}_{1i}(t) = x_{2i}(t) \ \forall i = 1, 2, \ldots, N \tag{3.14a}$$

$$\dot{x}_{2i}(t) = -x_{1i}(t) - \frac{1}{4}x_{2i}(t) - \frac{\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N} x_{1j}(t)x_{1k}(t)x_{1l}(t)\langle\phi_j(\Delta)\phi_k(\Delta)\phi_l(\Delta), \phi_i(\Delta)\rangle}{\langle\phi_i(\Delta), \phi_i(\Delta)\rangle}$$
$$\forall i = 1, 2, \ldots, N. \tag{3.14b}$$

Equation (3.14a) & Eqn. (3.14b) constitute the gPC dynamical system we used during propagation.

For the Duffing oscillator in Eqn. (3.10) & Eqn. (3.11), we assume that the initial states follows a standard Normal distribution ($\mu(t=0) = [0, 0],\ \Sigma(t=0) = \text{diag}(1, 1)$). We plot the mean and variance of the states obtained from gPC and from linearized dynamics. The state transition matrix of linearized system is given by

$$A = \begin{bmatrix} 0 & 1 \\ -1 - 3x_2^2 & -1/4 \end{bmatrix}.$$

The linearized mean (μ) and covariance (Σ) propagation equations are given by,

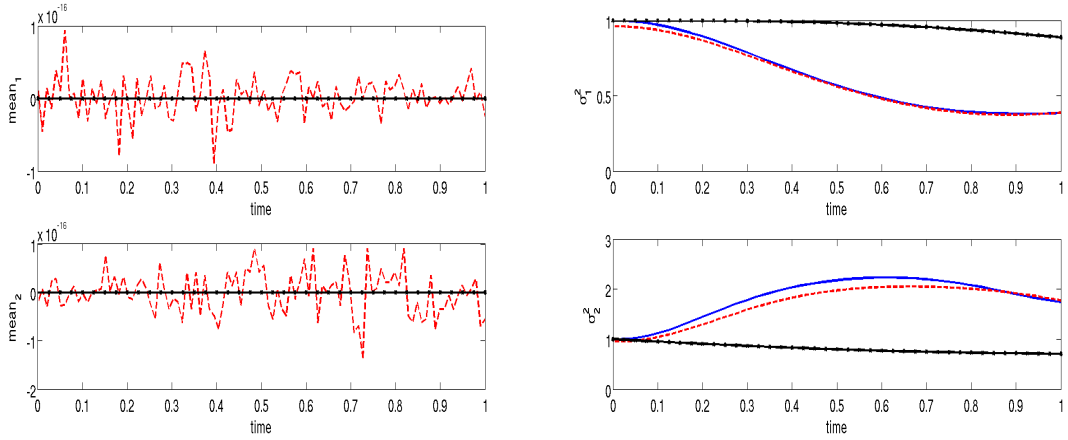
$$\dot{\mu}(t) = A\mu(t) \tag{3.15a}$$

$$\dot{\Sigma}(t) = A\Sigma(t) + \Sigma(t)A, \tag{3.15b}$$

where

$$\mu = \begin{bmatrix} \mathbb{E}[x_1] \\ \mathbb{E}[x_2] \end{bmatrix} \quad \Sigma = \begin{bmatrix} \mathbb{E}[x_1^2] - \mathbb{E}[x_1]^2 & \mathbb{E}[x_1x_2] \\ \mathbb{E}[x_1x_2] & \mathbb{E}[x_2^2] - \mathbb{E}[x_2]^2 \end{bmatrix}.$$

We compare the results obtained after application of gPC and propagation of linearized dynamics, to Monte Carlo (MC) simulations. Figure 4 shows the plots for



(a) Mean. (b) Variance.

Fig. 4.: Plots for a) mean b) variance of the Duffing oscillator in Eqn. (3.10) & Eqn. (3.11) obtained from gPC scheme (solid) and Monte Carlo (dashed) and linearized dynamics (star-solid) simulations.

evolution of mean and covariance. It can be seen that results for gPC and MC match closely, whereas the for the linearized dynamics, results diverge from that obtained from MC simulations. Hence, it can be concluded that gPC scheme effectively captures the evolution of uncertainty in the given nonlinear system.

C. Nonlinear State Estimation Using Polynomial Chaos and Higher Order Moments Update

Let us consider a nonlinear dynamical system being measured by a nonlinear measurement model. The states are given by $x \in \mathbb{R}^n$ and the measured outputs are, $\tilde{y} \in \mathbb{R}^m$. The dynamics is governed by,

$$\dot{x} = f(x, \Delta), \tag{3.16a}$$

$$\tilde{y} = h(x) + \nu, \tag{3.16b}$$

where ν is the measurement noise with $\nu \sim \mathcal{N}(0, R)$. The random parameters can be written as $\Delta = [\Delta_{x_0}\ \Delta_\rho]^T$, where Δ_{x_0} represents initial state uncertainty and Δ_ρ represents uncertainty in system parameters. Let $p(\Delta)$ be the distribution of Δ.

Estimation algorithms have essentially two steps, the propagation phase and the update phase. It is assumed that $p(\Delta)$ is stationary *during* the propagation phase. However, the distribution of $x(t, \Delta)$ will not be stationary due to the dynamics. Therefore, the distribution of Δ_{x_0} will change after every update phase. The distribution of Δ_ρ will typically not change at all, unless updated externally. Without loss of generality, here will will consider that the system has only initial state uncertainty. Let us also assume that the measurement updates are available at discrete time $t_k, t_{k+1}, \ldots$.

1. Step 1: Initialization of State

Given the probability density function of the parameters $p^k(\Delta)$ at time t_k, the initial condition for $X_{pc}(t_k)$ at t_k can be obtained using the following equation

$$x_{ij}(t_k) = \int_{\mathbf{D}_\Delta} \Delta_{x_{0_i}} \phi_j(\Delta) p^k(\Delta) d\Delta \text{ for } i = 1, \cdots, n;\ j = 0, \cdots, N, \tag{3.17}$$

where $\Delta_{x_0} = x(t_k, \Delta)$. The symbol $\Delta_{x_{0_i}}$ represents the i^{th} component of Δ_{x_0}, which is the random variable associated with initial condition uncertainty.

2. Step 2: Propagation of Uncertainty and Computation of Prior Moments

With initial condition defined by Eqn. (3.17), the system in Eqn. (3.7) is integrated over the interval $[t_k, t_{k+1}]$ to obtain $X_{pc}(t_{k+1})$, i.e.

$$X_{pc}(t_{k+1}) = X_{pc}(t_k) + \int_{t_k}^{t_{k+1}} F_{pc}(X_{pc}(\tau)) d\tau. \tag{3.18}$$

The moments of the random process $x(t,\Delta)$ at $t = t_{k+1}$ can be computed from $X_{pc}(t_{k+1})$ as follows. An outline of the derivation of the moments can be found in the previous section.

$$M_i^{1-} = x_{i0}, \tag{3.19a}$$

$$M_{ij}^{2-} = \sum_{p=0}^{N}\sum_{q=0}^{N} x_{ip}x_{jq}\langle\phi_p\phi_q\rangle, \tag{3.19b}$$

$$M_{ijk}^{3-} = \sum_{p=0}^{N}\sum_{q=0}^{N}\sum_{r=0}^{N} x_{ip}x_{jq}x_{kr}\langle\phi_p\phi_q\phi_r\rangle, \tag{3.19c}$$

$$M_{ijkl}^{4-} = \sum_{p=0}^{N}\sum_{q=0}^{N}\sum_{r=0}^{N}\sum_{s=0}^{N} x_{ip}x_{jq}x_{kr}x_{ls}\langle\phi_p\phi_q\phi_r\phi_s\rangle, \tag{3.19d}$$

and so on; for $i,j,k,l = 1,\cdots,n$. In the above equations $x_{ij} := x_{ij}(t_{k+1})$ and M^{i-} represents the i^{th} prior moment at t_{k+1}. The inner products of the basis functions are computed with respect to $p^k(\Delta)$, i.e.

$$\langle\phi_p\phi_q\phi_r\phi_s\rangle = \int_{\mathbf{D}_\Delta} \phi_p(\Delta)\phi_q(\Delta)\phi_r(\Delta)\phi_s(\Delta)p^k(\Delta)d\Delta.$$

3. Step 3: Update Phase

We incorporate the measurements $\tilde{y} := \tilde{y}(t_k)$ and the prior moments M^{i-} to get the posterior estimates of the moments, M^{i+}. Here we consider the prior state estimate $\hat{x}^-$ to be the expected value of $x(t,\Delta)$ at t_{k+1}, i.e. $\hat{x}^- = M^{1-} = \mathbb{E}[x]$. Also, $\hat{y}^- = h(\hat{x}^-)$. Let

$$v = \tilde{y} - h(\hat{x}^-) = h(x) + \nu - h(\hat{x}^-).$$

Using the approach used by Julier *et al.* [89, 90] and Park *et al.* [91], we use a linear Kalman gain K to update the moments. Although updates with nonlinear gains are also possible, but they were not considered in this dissertation. Linear update law has also been used by Majji *et al.* [92] in their design of nonlinear estimators using

higher order moment updates. The Kalman gain K is computed as

$$K = P^{xv}(P^{vv})^{-1}, \tag{3.20}$$

where $P^{xv}_{ij} = \mathbb{E}\left[x_i v_j^T\right]$ and $P^{vv}_{ij} = \mathbb{E}\left[v_i v_j^T\right]$. This gain is optimal in the minimum variance sense. The update equations for the moments are therefore given by,

$$M^{1+} = M^{1-} + Kv \tag{3.21a}$$

$$M^{2+} = M^{2-} - KP^{vv}K^T \tag{3.21b}$$

$$M^{3+} = M^{3-} + 3K^2P^{xvv} - 3KP^{xxv} - K^3P^{vvv} \tag{3.21c}$$

$$M^{4+} = M^{4-} - 4KP^{xxxv} + 6K^2P^{xxvv} - 4K^3P^{vvvx} + K^4P^{vvvv}. \tag{3.21d}$$

The work of Majji *et al.* [92] and Julier *et al.* [89], contains complete derivations of these formula. The fifth and higher order moment update equations can be computed in a similar fashion, using appropriate prior moments and Kalman gain [89, 92].

The tensors $P^{xv}, P^{vv}, P^{xvv}, P^{xxv}, P^{vvv}, P^{xxvx}, P^{xxvv}$, P^{vvvx}, and P^{vvvv} can be computed in terms of the gPC coefficients. Here we only the derivations for P^{xv} and P^{vv}, the other tensors can be calculated in a similar manner. Using this derivation process, the higher order tensors can be computed very easily. The expressions for P^{xv} and P^{vv} are

$$\begin{aligned}
P^{xv} = & \quad \mathbb{E}\left[(x-\hat{x})(\tilde{y}-\hat{y})^T\right] = \mathbb{E}\left[x\tilde{y}^T\right] - \hat{x}\mathbb{E}\left[\tilde{y}^T\right] \\
= & \quad \mathbb{E}\left[x(h(x)+\nu)^T\right] - \hat{x}\mathbb{E}\left[(h(x)+\nu)^T\right] \\
= & \quad \mathbb{E}\left[xh^T(x)\right] - \hat{x}\mathbb{E}\left[h^T(x)\right] \\
= & \quad \int_{\mathbf{D}_\Delta} \left(\sum_i \mathbf{x}_i\phi_i(\Delta)\right) h^T \left(\sum_j \mathbf{x}_j\phi_j\right) p^k(\Delta)d\Delta \\
& \quad -\hat{x}\int_{\mathbf{D}_\Delta} h^T\left(\sum_j \mathbf{x}_j\phi_j(\Delta)\right) p^k(\Delta)d\Delta,
\end{aligned} \tag{3.22}$$

where $\mathbf{x}_i = [x_{1i}\ x_{2i}\ \cdots\ x_{ni}]^T$. Similarly,

$$P^{vv} = \int_{\mathbf{D}_\Delta} h\left(\sum_j \mathbf{x}_j \phi_j(\Delta)\right) h^T\left(\sum_j \mathbf{x}_j \phi_j(\Delta)\right) p^k(\Delta) d\Delta + R - \hat{y}\hat{y}^T. \tag{3.23}$$

The above expressions for P^{xv} and P^{vv} are dependent on the output function $h(x)$. If $h(x)$ is polynomial function, the above expressions become functions of M^{i-}. This is shown through an example in a later section, where we consider the output equation $h(x) = x^T x$. When $h(x)$ is a transcendental function, computation of the above integrals cannot be performed directly. One approach would be to expand $h(x)$ about $\mathbb{E}[x]$ in terms of the perturbations, using Taylor series expansion [87], and obtain a polynomial approximation of $h(x)$. While Taylor series approximation is straightforward and generally computationally cost effective, it becomes severely inaccurate when higher order gPC expansions are required to represent the physical variability [87]. For example, a 5^{th} order Taylor series approximation using 3^{rd} order gPC expansion would require tensor products of six 3^{rd} order basis functions. This will result in 18^{th} order polynomials. This will increase if higher order Taylor series or gPC expansions are used to obtain better approximations. It is well known that computation of higher order polynomials using finite significant digits representation of real numbers have associated numerical errors. At the same time, for many non-linear functions this Taylor series approximation is limited by the theoretical range of convergence Taylor series. To tackle the problem of inaccuracies in the evaluation of transcendental functions, using Taylor series expansions, a more robust algorithm is presented by Debusschere *et al.* [87]. This method is valid for any non polynomial function $u(\xi)$ for which $du/d\xi$ can be expressed as a rational function of $\xi, u(\xi)$. The same issues are encountered with $f(x, \Delta)$, which defines the dynamics of the system.

4. Step 4: Estimation of Posterior Probability Distribution

After the update step, we only have the moments of the of the posterior PDF, $p^{k+1}(\Delta)$ at time t_{k+1}. Hence it is difficult to estimate $p^{k+1}(\Delta)$, except when Gaussian behavior is assumed [93]. In this estimation algorithm, the posterior PDF, is determined using maximum entropy estimation theory, subject to constraints defined by the posterior moments M^{i-}. This is the solution of the following optimization problem,

$$J := \max_{p^{k+1}(\Delta)} \; -\int_{D_\Delta} p^{k+1}(\Delta)\log(p^{k+1}(\Delta))\,d\Delta, \tag{3.24}$$

subject to

$$C_1 := \int_{D_\Delta} \Delta p^{k+1}(\Delta)\,d\Delta = M^{1+}, \tag{3.25a}$$

$$C_2 := \int_{D_\Delta} Q_2(\Delta) p^{k+1}(\Delta)\,d\Delta = M^{2+}, \tag{3.25b}$$

$$C_3 := \int_{D_\Delta} Q_3(\Delta) p^{k+1}(\Delta)\,d\Delta = M^{3+}, \tag{3.25c}$$

and so on. The symbols $Q_i(\Delta)$ are tensors of polynomials defining the moments corresponding to M^{i+}. The functional space of $p^{k+1}(\Delta)$ is approximated using Gaussian mixture models (GMM) [80]. GMM are dense in the space of continuous functions and a sufficiently large mixture can exactly approximate $p^{k+1}(\Delta)$ [94]. In general, any parameterization of $p^{k+1}(\Delta)$ is possible, which is known under the name of kernel density estimation [95]. Under GMM approximation, $p^{k+1}(\Delta)$ is parameterized as

$$p^{k+1}(\Delta) = \sum_{i=1}^{M} \alpha_i \mathcal{N}(\mu_i, \Sigma_i), \tag{3.26}$$

with $\alpha_i \in \mathbb{R}$, $\mu_i \in \mathbb{R}^n$, $\Sigma_i = \Sigma_i^T \in \mathbb{R}^{n \times n}$ and $\Sigma_i \geq 0$. For computational simplicity, Σ_i is assumed to be diagonal, i.e. $\Sigma_i = \text{diag}(\sigma_{i1} \ \cdots \ \sigma_{in})$. For $p^{k+1}(\Delta)$ to be a PDF the following constraints need to be satisfied:

$$\int_{\mathcal{D}_\Delta} p^{k+1}(\Delta) d\Delta = 1 \Rightarrow \sum_{i}^{M} \alpha_i = 1, \tag{3.27}$$

$$p^{k+1}(\Delta) \geq 0 \Rightarrow \alpha_i \geq 0. \tag{3.28}$$

Equation (3.27) and Eqn. (3.28) together implies

$$0 \leq \alpha_i \leq 1. \tag{3.29}$$

With GMM approximation of $p^{k+1}(\Delta)$, the integrals in Eqn. (3.24) and Eqn. (3.25a) through Eqn. (3.25c) can be analytically computed and expressed in terms of the unknowns α_i, μ_i and Σ_i. For a given term in the summation in Eqn. (3.26), and for $x \in \mathbb{R}^3$, the cost function J_i and the constraints $C_{i,j}$, are given by the following expressions.

$$J_i = \alpha_i(\log(2) - \log(\alpha_i) + \log(\sigma_{i1}) + \log(\sigma_{i2}) + \log(\sigma_{i3}) + \log(\pi) + 1)$$

$$C_{i,1} = \mu_{i1}\alpha_i,$$

$$C_{i,2} = \mu_{i2}\alpha_i,$$

$$C_{i,3} = \mu_{i3}\alpha_i,$$

$$C_{i,4} = \mu_{i1}^2\alpha_i + \alpha_i\sigma_{i1}^2,$$

$$C_{i,5} = \mu_{i1}\mu_{i2}\alpha_i,$$

$$C_{i,6} = \mu_{i2}^2\alpha_i + \alpha_i\sigma_{i2}^2,$$

$$C_{i,7} = \mu_{i1}\mu_{i3}\alpha_i,$$

$$C_{i,8} = \mu_{i2}\mu_{i3}\alpha_i,$$

$$C_{i,9} = \mu_{i3}^2\alpha_i + \alpha_i\sigma_{i3}^2,$$

$$C_{i,10} = \mu_{i1}^3\alpha_i + 3\mu_{i1}\alpha_i\sigma_{i1}^2,$$

$$C_{i,11} = \mu_{i1}^2\mu_{i2}\alpha_i + \mu_{i2}\alpha_i\sigma_{i1}^2,$$

$$C_{i,12} = \mu_{i1}\mu_{i2}^2\alpha_i + \mu_{i1}\alpha_i\sigma_{i2}^2,$$

$$C_{i,13} = \mu_{i2}^3\alpha_i + 3\mu_{i2}\alpha_i\sigma_{i2}^2,$$

$$C_{i,14} = \mu_{i3}^3\alpha_i + 3\mu_{i3}\alpha_i\sigma_{i3}^2,$$

$$C_{i,15} = \mu_{i1}\mu_{i3}^2\alpha_i + \mu_{i1}\alpha_i\sigma_{i3}^2,$$

$$C_{i,16} = \mu_{i2}\mu_{i3}^2\alpha_i + \mu_{i2}\alpha_i\sigma_{i3}^2,$$

$$C_{i,17} = \mu_{i1}^2\mu_{i3}\alpha_i + \mu_{i3}\alpha_i\sigma_{i1}^2,$$

$$C_{i,18} = \mu_{i3}\mu_{i2}^2\alpha_i + \mu_{i3}\alpha_i\sigma_{i2}^2,$$

$$C_{i,19} = \mu_{i3}\mu_{i2}\mu_{i1}\alpha$$

The expressions in the above equations can also be obtained analytically for $x \in \mathbb{R}^n$. The number of constraints $C_{i,j}$ needed for $x \in \mathbb{R}^n$, is given by the following

expression.

$$\dim(C_{i,j}) = \sum_{k=1}^{3} \binom{n+k-1}{k} \; \forall i = 1, \cdots, M \tag{3.30}$$

Detailed derivation of the above formula has been omitted here, and can be found in the literature [96].

If the moments are represented in column form as,

$$M^{1+} = [M_1^{1+}, M_2^{1+}, M_3^{1+}]^T$$
$$M^{2+} = [M_{11}^{2+}, \cdots, M_{13}^{2+}, M_{21}^{2+}, \cdots, M_{23}^{2+}, \cdots]^T$$
$$M^{3+} = [M_{111}^{2+}, \cdots, M_{113}^{3+}, M_{121}^{3+}, \cdots, M_{123}^{3+}, M_{131}^{3+}, \cdots, M_{133}^{3+}, \cdots]^T.$$

The cost J in Eqn. (3.24) and constraints C_j in Eqn. (3.25a) through Eqn. (3.25c) are given by,

$$J = \sum_{i=1}^{M} J_i$$
$$C_1 := \sum_{i=1}^{M} C_{i,j} = M_j^{1+} \; j = 1, \cdots, 3$$
$$C_2 := \sum_{i=1}^{M} C_{i,j} = M_{j-3}^{2+} \; j = 4, \cdots, 9$$
$$C_3 := \sum_{i=1}^{M} C_{i,j} = M_{j-9}^{3+} \; j = 10, \cdots, 19.$$

It can be seen that the cost function is convex in α_i and concave in σ_{ij}. The constraints are convex in α_i and σ_{ij}, and not all the constraints are convex in μ_{ij}. The problem can be made convex by restricting $\mu_{ij} \geq 0$, which will require affine transformation of the state variables and rewriting the dynamics and output equation in terms of the new variables. The optimization problem, as presented here, can also be solved as a nonlinear programming problem.

5. Step 5: Generation of Basis Functions

Once $p^{k+1}(\Delta)$ has been estimated using GMM approximations, the basis functions $\{\phi_i(\Delta)\}$ need to be generated so that they are orthogonal with respect to the new probability density function. If $\{\phi_i(\Delta)\}$ are orthogonal, the gPC approximations have exponential convergence and thus are optimal. For any other basis functions, the approximation is worse than optimal. However, for some applications use of the same basis functions for all the time steps may provide acceptable results. We use the Gram-Schmidt procedure to generate the set of basis functions $\{\phi_i(\Delta)\}$ that are orthogonal with respect to $p^{k+1}(\Delta)$. Consequently, all the inner products, in the estimation algorithm presented, need to be recomputed at every time step.

The estimation algorithm proceeds by repeating steps one to five as described above.

D. Nonlinear Estimation Using Polynomial Chaos and Update Using Bayesian Inference

The estimation algorithm presented in th previous section was based on updating higher order moments and estimating the posterior PDF. Since, the prior moments is available, we can directly estimate the prior PDF from moments and then use Bayesian inference to get the posterior PDF.

In section we use the Bayesian inference to develop the nonlinear estimation algorithm using polynomial chaos theory. At time t_{k+1}, let $x_{k+1}(\Delta)$ be the random variable associated with the state, let $\hat{x}_{k+1}$ be the state estimate and $\tilde{y}_{k+1}$ be the measurement. The objective is to incorporate $\tilde{y}_{k+1}$ to determine $\hat{x}_{k+1}$, using the classical Bayesian approach (pg. 377, in ref.[97]). In this framework we first calculate

$p(\tilde{y}_{k+1}|x_{k+1}(\Delta) = \hat{x}_{k+1})$, the conditional probability density function of $\tilde{y}_{k+1}$, given the current state estimate $\hat{x}_{k+1}$. This function is the likelihood function. Next,we find the conditional PDF of the state, given $\tilde{y}_{k+1}$, i.e. $p(x_{k+1}(\Delta)|\tilde{y}_{k+1})$, which is the posterior PDF, from the prior PDF and the likelihood function. The state estimate $\hat{x}_{k+1}$ is then determined using the maximum likelihood, minimum-variance or minimum error criterion from $p(x_{k+1}(\Delta)|\tilde{y}_{k+1})$. The algorithm in the context of polynomial chaos is described next.

1. Step 1: Initialization of State and Propagation of Uncertainty

The first two steps in the Bayesian estimation algorithm using polynomial chaos theory is the same as the previous algorithm. Given the PDF $p^k(\Delta)$, the initial states for $X_{pc}(t_k)$ are determined using Eqn. (3.17) . The propagation of uncertainty is achieved by integrating the system in Eqn. (3.7) over the interval $[t_k, t_{k+1}]$ to obtain $X_{pc}(t_{k+1})$. With $X_{pc}(t_{k+1})$, higher order moments are obtained as described in the previous algorithm. With these moments, the prior probability density function $p^{k+1^-}(\Delta)$ is determined using the Gaussian mixture model, also as described in the previous algorithm. Polynomial chaos is used only to determine $p^{k+1^-}(\Delta)$. The following steps are standard steps in Bayesian estimation and have been presented here for completeness.

2. Step 2: Calculating the Posterior Probability Density Function

First, the likelihood function $p(\tilde{y}_{k+1}|x_{k+1}(\Delta) = \hat{x}_{k+1})$, is constructed assuming Gaussian measurement noise and the sensor model as shown in Eqn. (3.16a) & Eqn. (3.16b). It is defined as

$$p(\tilde{y}_{k+1}|x_{k+1}(\Delta) = \hat{x}_{k+1}) = \frac{1}{\sqrt{(2\pi)^m|R|}} e^{-\frac{1}{2}(\tilde{y}_{k+1}-h(\hat{x}_{k+1})^T R^{-1}(\tilde{y}_{k+1}-h(\hat{x}_{k+1})}, \qquad (3.31)$$

where $|R|$ is the determinant of measurement noise covariance matrix.

The posterior probability density function of the states is determined next. It is given by the density function of the states given the current measurement, $p(x_{k+1}(\Delta)|\tilde{y}_{k+1})$. Using the classical Bayes rule we can write it as,

$$p(x_{k+1}(\Delta)|\tilde{y}_{k+1}) = \frac{p(\tilde{y}_{k+1}|x_{k+1}(\Delta) = \hat{x}_{k+1})p^{k+1^-}(\Delta)}{\int_{\mathbf{D}_\Delta} p(\tilde{y}_{k+1}|x_{k+1}(\Delta) = \hat{x}_{k+1})p^{k+1^-}(\Delta)d\Delta}, \tag{3.32}$$

where $p^{k+1^-}(\Delta)$ is the prior PDF.

3. Step 3: Getting the State Estimate

Depending on the criterion function for state estimation, we can compute the estimate $\hat{x}_{k+1}$ from $p(x_{k+1}(\Delta)|\tilde{y}_{k+1})$ The commonly used criterion are:

1. *Maximum likelihood estimate*: maximize the probability that $\hat{x}_{k+1} = x_{k+1}(\Delta)$. This translates to,

$$\hat{x}_{k+1} = \textbf{mode of } p(x_{k+1}(\Delta)|\tilde{y}_{k+1}).$$

2. *Minimum variance estimate*, which translates to,

$$\hat{x}_{k+1} = \min_{x_{k+1}(\Delta)} \int_{\mathbf{D}_\Delta} ||x_{k+1}(\Delta) - \hat{x}_{k+1}| \, |p(x_{k+1}(\Delta)|\tilde{y}_{k+1})d\Delta = \mathbb{E}\left[x_{k+1}(\Delta)\right].$$

3. *Minimum error estimate*, which translates to,

$$\hat{x}_{k+1} = \min_{x_{k+1}(\Delta)} \sup |x_{k+1}(\Delta) - \hat{x}_{k+1}| = \textbf{median of } p(x_{k+1}(\Delta)|\tilde{y}_{k+1}).$$

4. Step 4: Regeneration of Basis Functions

With $p^{k+1}(\Delta) := p(x_{k+1}(\Delta)|\tilde{y}_{k+1})$, the basis functions $\{\phi_i(\Delta)\}$ should be generated so that they are orthogonal with respect to the new PDF. This step is identical to Step 5, in the previous algorithm. The estimation algorithm proceeds by repeating steps one to four as described above.

Here, two algorithms have been proposed based on gPC, which differ in their update methodology. Flowcharts of the PC based estimation algorithms are shown in Fig. 5.

E. Application to Duffing Oscillator

We apply the estimation algorithm presented here to the classical Duffing oscillator. It is a two state system, $x = [x_1, x_2]^T$, whose dynamical equations are given by Eqn. (3.10) & Eqn. (3.11). After application of gPC scheme on the dynamical system, the gPC dynamical system is given by Eqn. (3.14a) & Eqn. (3.14b).

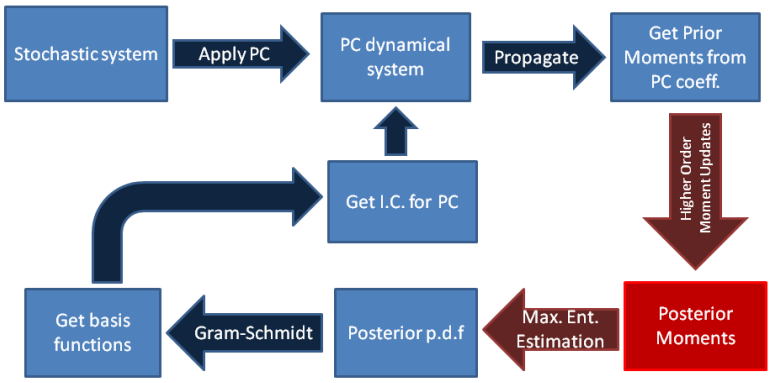
In this example, we consider initial condition uncertainty in x with Gaussian distribution, i.e. $\Delta_{x_0} \sim \mathcal{N}(\mu, \Sigma)$. We use a scalar measurement model for the Duffing oscillator which is given by

$$\tilde{y} = x^T x + \nu, \tag{3.33}$$

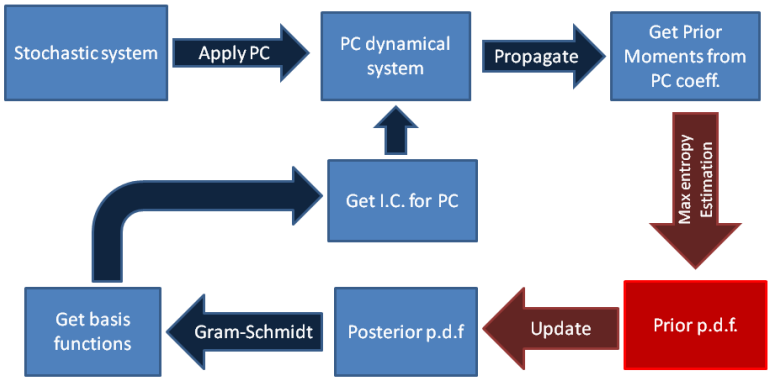
with $\mathbb{E}[\nu] = 0$ and $\mathbb{E}\left[\nu\nu^T\right] = 0.006$. First, we will show results when we apply higher order moments update. Results for the estimation algorithm using Bayesian inference will be discussed next.

1. Results for Estimation with Higher Order Moments Update

In this example we update up to third order moments. Therefore, we only need to compute the tensors $P^{xv}, P^{vv}, P^{xvv}, P^{xxv}, P^{vvv}$, which for this system are given in



(a) Estimation with higher order moments update.



(b) Estimation using Bayesian inference.

Fig. 5.: Flowchart of polynomial chaos based estimation algorithms.

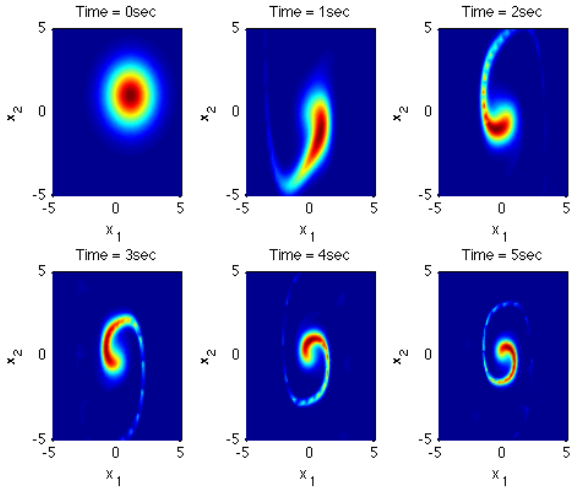


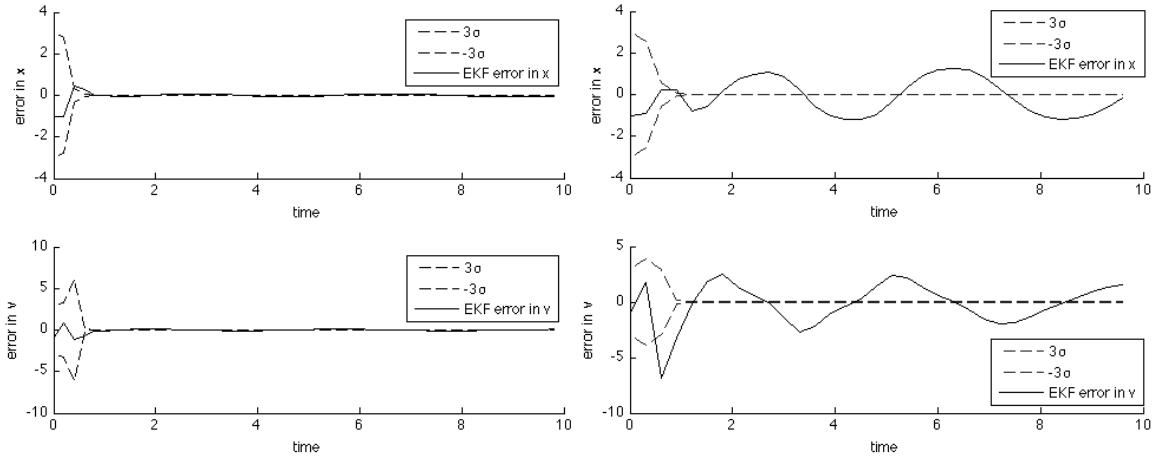
Fig. 6.: Evolution initial condition uncertainty for Duffing oscillator. Probability density functions were obtained using high fidelity Monte Carlo simulations.

Appendix C, Eqn. (C.1).

The prior moments M^{i-}, for $i = 1, \cdots, 6$ are computed from the gPC coefficients X_{pc} and the probability density function $p^k(\Delta)$. Using the tensors P^{xv}, P^{vv}, P^{xvv}, P^{xxv}, P^{vvv}, we obtain the posterior moments M^{i+}, for $i = 1, 2, 3$, and subsequently obtain the estimate of the PDF for the next time step, i.e. for $p^{k+1}(\Delta)$. In this example, we solved the associate optimization problem as a nonlinear programming problem, using SNOPT [98]. In our future work we will solve this problem in the convex optimization framework.

Figure 6 shows the evolution of initial condition uncertainty for the Duffing oscillator. We can observe that the probability density function for the state evolve in a non Gaussian manner. Intuitively, we can say that Gaussian assumption of PDF will not perform well here.

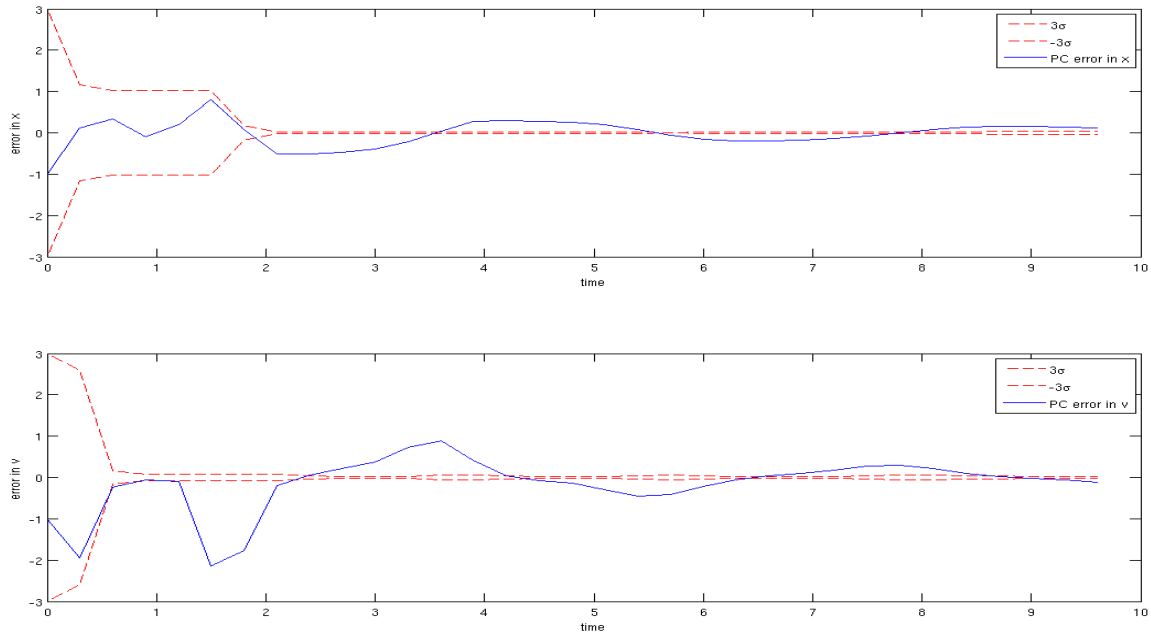
Figure 7 and Fig. 8 illustrate the performance of the EKF based estimator and the proposed gPC based estimator. The true initial condition for the system is taken to be $[2\ 2]^T$ and the uncertainty is assumed to be Gaussian with $\mu = [1\ 1]^T$ and



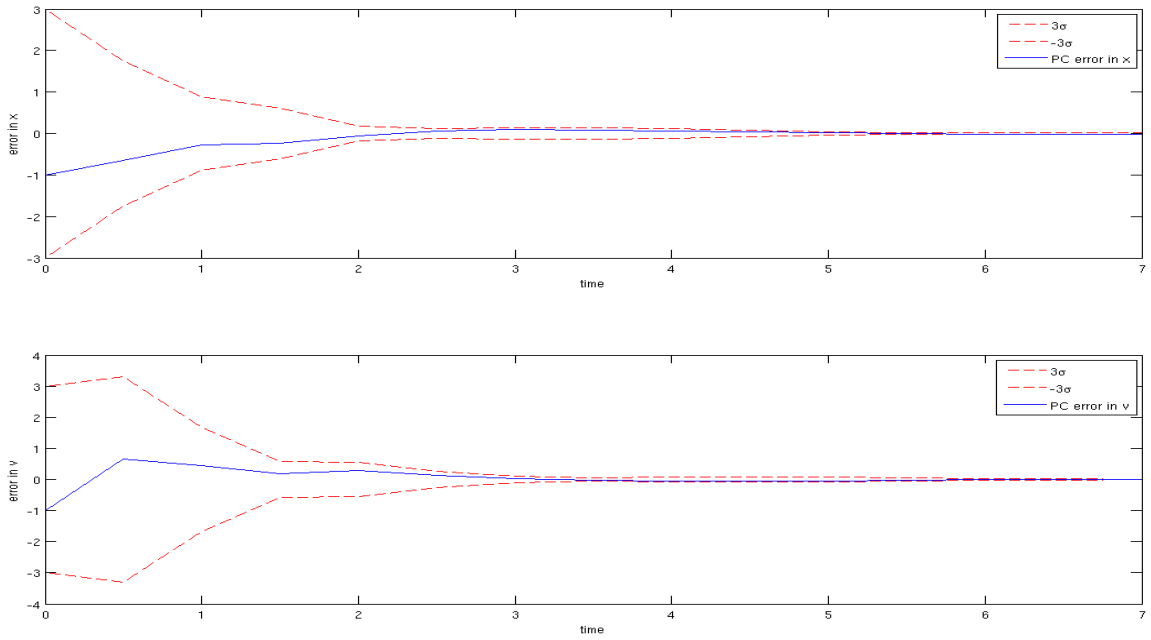
(a) EKF based estimator with 0.2s update. (b) EKF based estimator with 0.3s update.

Fig. 7.: Performance of EKF estimators. Dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

$\Sigma = \text{diag}(1, 1)$. Therefore, the initial error in the state estimation is $[1\ 1]^T$. Figure 7a shows the performance of the EKF based estimator with 0.2 seconds measurement update. We observe that the errors in estimates go to zero rapidly. However, when the measurement update interval is increased to 0.3 seconds, the EKF based estimator performs poorly and the errors do not converge to zero. Figure 8a shows the plots of the estimator when gPC theory is used to propagate the uncertainty and only the first two moments are updated, using standard Kalman update law. We observe that this combination achieves better results than the EKF estimator but is inconsistent as the errors escape the $\pm 3\sigma$ bounds. Figure 8b shows the performance of the gPC based estimator with third order moment updates, performed every 0.5 seconds. We observe that errors in the estimates converge to zero rapidly and are within the $\pm 3\sigma$ bounds. The EKF based estimator for this case diverges. This highlights the importance of using nonlinear dynamics for uncertainty propagation along with higher order moments update for solving nonlinear estimation problems.



(a) gPC based estimator with first two moments updated every 0.3s.



(b) gPC based estimator with first three moments updated every 0.5s.

Fig. 8.: Performance of gPC estimators. Dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

2. Results for Estimation Using Bayesian Inference

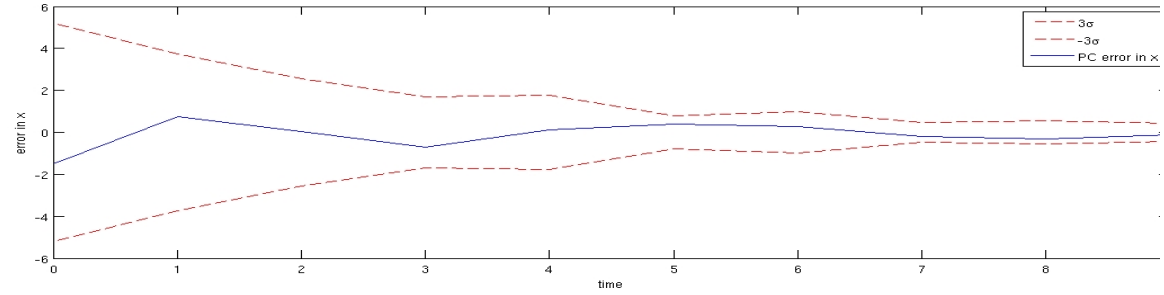
In this section we discuss the results of the application of estimation algorithm that uses Bayesian inference for update, to the Duffing oscillator in Eqn. (3.10) & Eqn. (3.11). The true initial states of the system were taken to be $[2.5, 2.5]$. The initial state uncertainty was assumed to be Gaussian with mean and covariance matrix given by

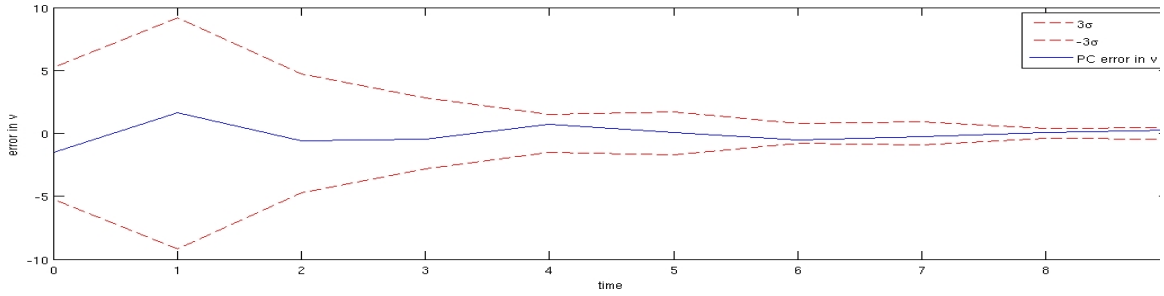
$$\mu = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \Sigma^2 = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}.$$

Hence, the initial error in estimation and the initial state uncertainty is larger than the previous case. Time between each measurement update was taken to be 1s. Figure 9 shows plots for estimation error and $\pm 3\sigma$ bounds for minimum covariance, maximum likelihood and minimum error estimators, and EKF based estimator. It can be seen that for the PC based estimators, errors converge and are within the $\pm 3\sigma$ limits, whereas for EKF errors diverge and escape outside the $\pm 3\sigma$ bounds.

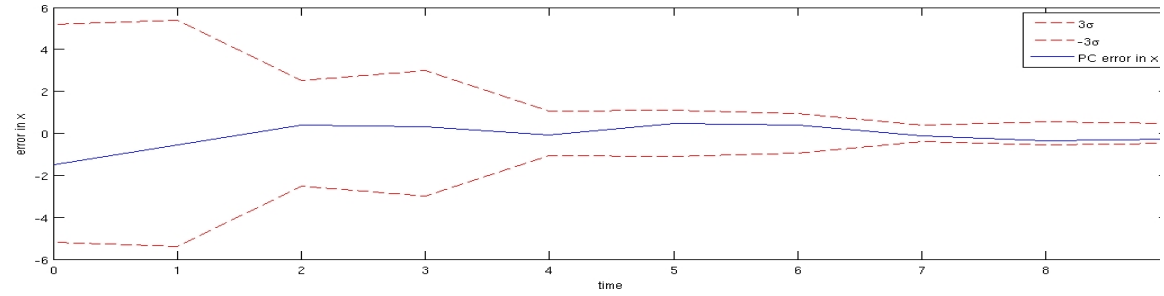
F. Application to Hypersonic Reentry

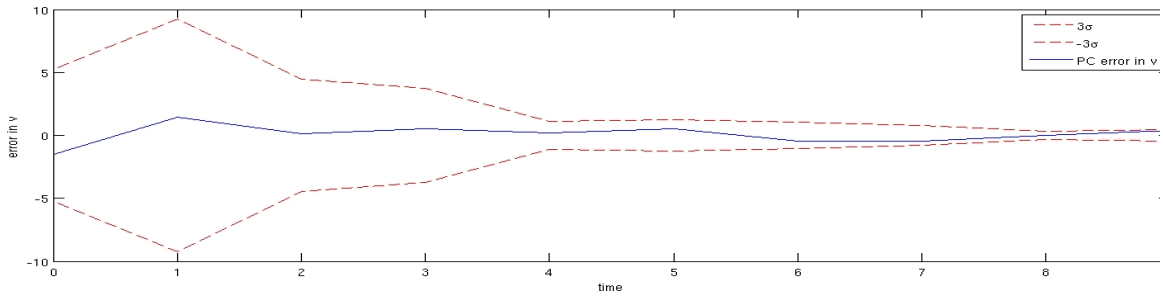
We apply the estimation algorithm, that uses Bayesian inference for PDF update, to reentry of a hypersonic vehicle in the atmosphere of Mars. We assume only longitudinal motion of the reentry vehicle here. The simplified dynamics of reentry are represented by Vinh's equation [99], in three states - the distance from planet's center r, velocity v, and the flight-path angle γ, or $x = [r\ v\ \gamma]^T$. The equations can be





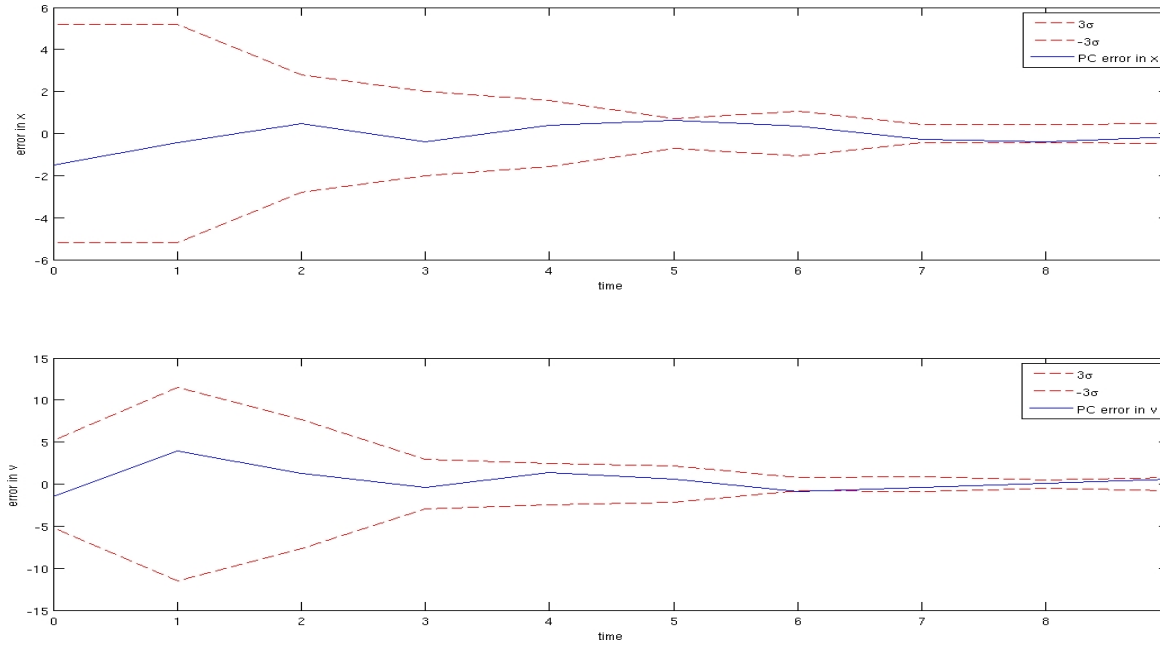
(a) Minimum covariance estimator.



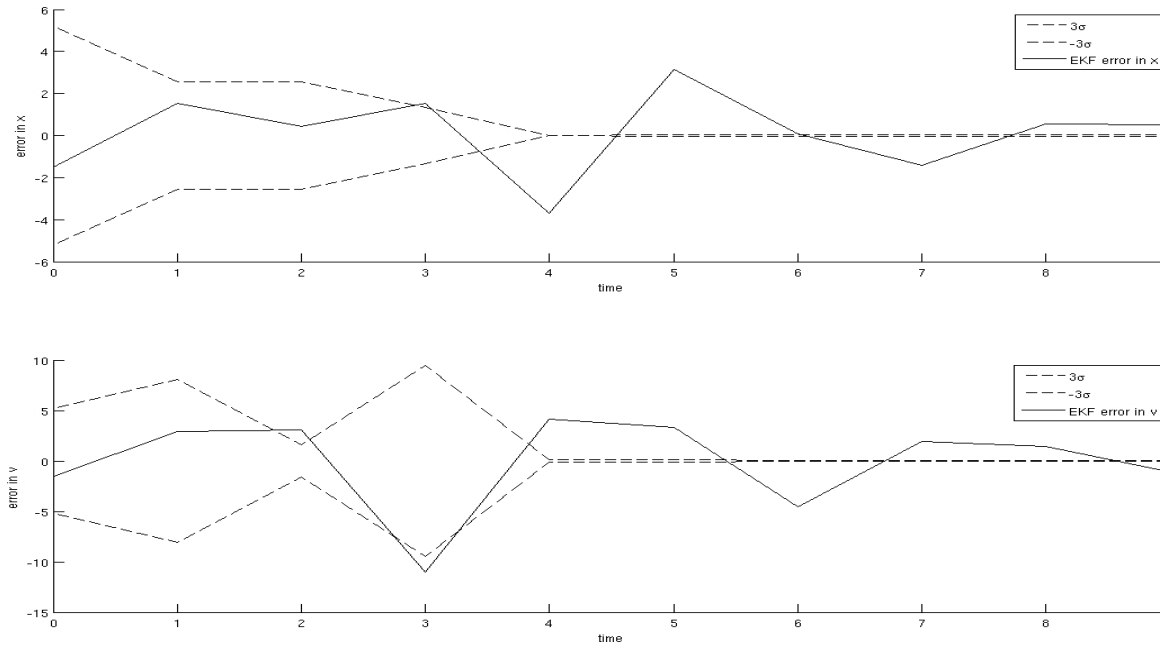


(b) Maximum likelihood estimator.

Fig. 9.: Plots for time (x-axis) vs. $\pm 3\sigma$ limits (dashed lines) and estimation error (solid lines) (y-axis) for gPC based estimators and EKF based estimator.



(c) Minimum error estimator.



(d) EKF based estimator.

Fig. 9.: Continued.

Table III.: Explanation and values of the constants for Martian atmosphere (Sengupta and Bhattacharya, 2008)

Description of Constants	**Value**
Radius of Mars	$R_m = 3397 \times 10^3$ m
Acceleration due to Gravity of Mars	$g = 3.71$ m/s^2
Ballistic Coefficient of the Vehicle	$B_c = 72.8$ kg/m^2
Lift-to-Drag Ratio of the Vehicle	$\frac{L}{D} = 0.3$
Density at the Surface of Mars	$\rho_0 = 0.0019$ kg/m^3
Scale Height for Density Computation	$h_1 = 9.8$ km $h_2 = 20$ km
Escape Velocity of Mars	$v_c = 5.027$ km/s

written as

$$\dot{r} = v\sin(\gamma) \tag{3.34a}$$

$$\dot{v} = -\frac{\rho v^2}{2B_c} - g\sin(\gamma) \tag{3.34b}$$

$$\dot{\gamma} = \left(\frac{v}{r} - \frac{g}{v}\right)\cos(\gamma) + \frac{\rho}{2B_c}\left(\frac{L}{D}\right)v, \tag{3.34c}$$

where g is the acceleration due to gravity, B_c is the ballistic coefficient, $\frac{L}{D}$ is the lift-to-drag ratio of the the vehicle, ρ is the atmospheric density given by

$$\rho = \rho_0 e^{(\frac{h_2 - h}{h_1})},$$

where ρ_0, h_1 and h_2 are constants depending on the planet's atmospheric model, $h = r - R_m$ is the height above planet's surface and R_m is the radius of the planet. Choices of the constants in Eqn. (3.34a) through Eqn. (3.34c) used to simulate reentry in Martian atmosphere are given in Table III [100].

The measurement model, $\tilde{y}$, consists of the dynamic pressure $\bar{q}$, the heating rate H [101], and the flight path angle, or $\tilde{y} = [\bar{q}\ H\ G]^T$, whose expressions are,

$$\bar{q} = \frac{1}{2}\rho v^2 \tag{3.35a}$$

$$H = k\rho^{\frac{1}{2}} v^{3.15} \tag{3.35b}$$

$$G = \gamma, \tag{3.35c}$$

where $k = 4.47228 \times 10^{-9}$ is the scaled material heating coefficient.

Here, initial state uncertainty with Gaussian distribution has been considered, i.e. $x(0, \Delta) = \Delta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ where, μ_0 and σ_0 are mean and standard deviation, respectively and have the values,

$$\mu_0 = [R_m + 54 \text{ km}, 2.4 \text{ km/s}, -9^o]^T \tag{3.36a}$$

$$\sigma_0 = \begin{pmatrix} 5.4 \text{ km} & 0 & 0 \\ 0 & 240 \text{ m/s} & 0 \\ 0 & 0 & 0.9^o \end{pmatrix}. \tag{3.36b}$$

To achieve consistency in dimensions, every constants in Eqn. (3.34a) through Eqn. (3.34c) are scaled appropriately, to create a non-dimensionalized system. The constants scaling the base units are given in Table IV.

Table IV.: Scaling constants for base units

Units	**Scaling Constants**
Mass	Mass of the vehicle = 2800 kg
Length	Radius of Mars (R_m)
Time	$\frac{\text{Radius of Mars}}{\text{Escape Velocity of Mars}} = 675.7$ s

The measurements are normalized to lie within $[-1, 1]$, so that they have consistent magnitude. The normalization factors used for measurements are given in Table V. The measurement noise ν is assumed to have mean and covariance as,

Table V.: Normalization factors for measurements

Measurement	**Normalization Factors**
Dynamic Pressure	$1.97 \times 10^3 \frac{\mathrm{N}}{\mathrm{m}}$
Heating Rate	$0.0231 \frac{\mathrm{J}}{\mathrm{m\text{-}s}}$
Flight Path Angle	19.13^o

$\mathbb{E}[\nu] = [0, 0, 0]^T$ and $R = \mathbb{E}\left[\nu\nu^T\right] = 6 \times 10^{-5}\mathcal{I}_3$.

As mentioned before, gPC theory works best, when the nonlinearities are in the form of polynomials [87, 102]. In this case, the trigonometric and exponential terms in Eqn. (3.34a) through Eqn. (3.34c) are approximated, with cubic polynomials in γ and h respectively. For example, let $S(\gamma)$ be a suitable approximation of $\sin(\gamma)$, then,

$$S(\gamma) = a_0 + a_1\gamma + a_2\gamma^2 + a_3\gamma^3,$$

and

$$S'(\gamma) = a_1 + 2a_2\gamma + 3a_3\gamma^2.$$

The coefficients of $S(\gamma)$ are obtained by equating the values of $\sin(\gamma)$ and its derivative, with $S(\gamma)$ and $S'(\gamma)$ respectively, at the boundaries of γ's domain. For the

present case, $\gamma \in [-90^o, 0]$, which yields a system of linear equations given by,

$$\begin{pmatrix} 1 & -1.57 & 2.47 & -3.88 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & -3.14 & 7.40 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

which can be solved for a_0, a_1, a_2 and a_3. The approximation of $\cos(\gamma)$ can be found in a similar manner.

The approximation of sine and cosine terms for the present case are,

$$\sin(\gamma) \approx S(\gamma) := \gamma + 0.0574\gamma^2 - 0.1107\gamma^3$$
$$\cos(\gamma) \approx C(\gamma) := 1 - 0.5792\gamma^2 - 0.1107\gamma^3.$$

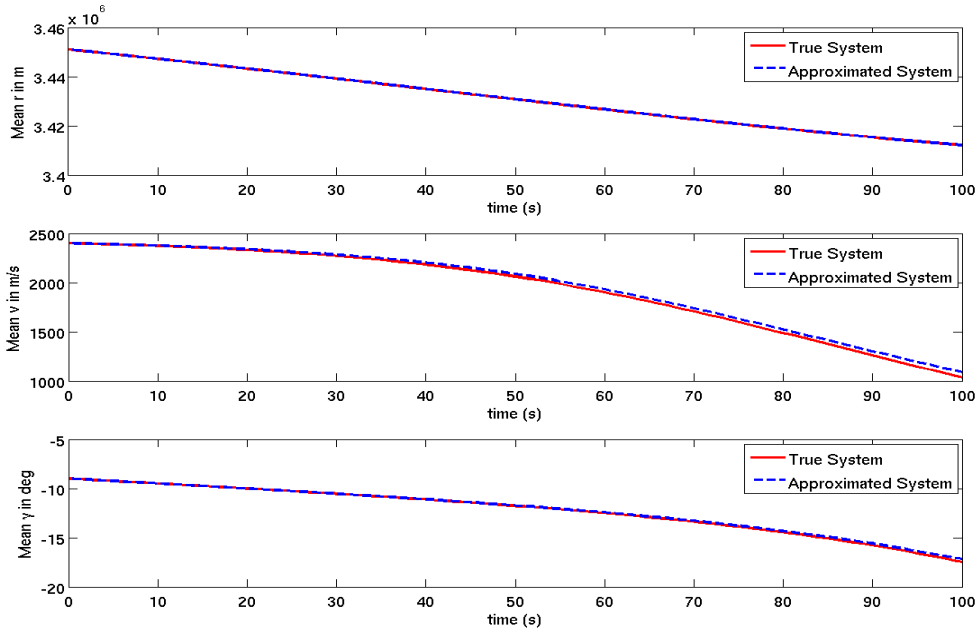
Assuming $h \in [0, 100 \text{ km}]$, the exponential density term, is approximated by cubic polynomials, in three different intervals. The approximated density term $D(h) \approx \rho$, is given by,

$$D(h) := \begin{cases} 0.0146 - 1.50 \times 10^{-6}h + 5.93 \times 10^{-11}h^2 - 8.45 \times 10^{-16}h^3 & h \in [0, 25\text{km}) \\ 0.0080 - 4.72 \times 10^{-7}h + 9.59 \times 10^{-12}h^2 - 6.61 \times 10^{-17}h^3 & h \in [25\text{km}, 50\text{km}) \\ 0.0015 - 5.2 \times 10^{-8}h + 5.90 \times 10^{-13}h^2 - 2.20 \times 10^{-18}h^3 & h \in [50\text{km}, 100\text{km}) \end{cases}.$$

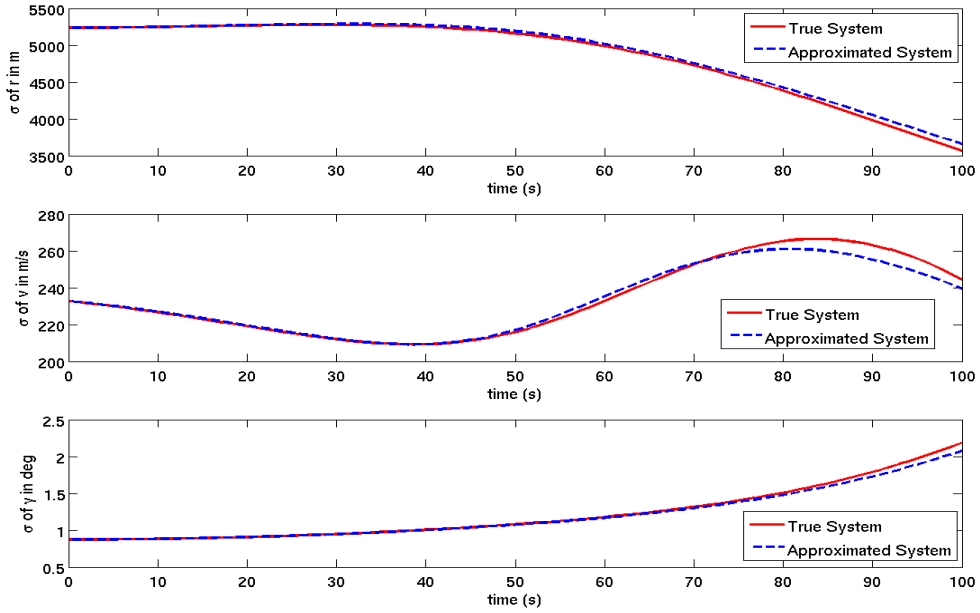
The approximated system can be written as

$$\dot{r} = vS(\gamma) \tag{3.37a}$$
$$\dot{v} = -\frac{D(h)v^2}{2B_c} - gS(\gamma) \tag{3.37b}$$
$$\dot{\gamma} = \left(\frac{v}{r} - \frac{g}{v}\right) C(\gamma) + \frac{1}{2B_c}\left(\frac{L}{D}\right) v. \tag{3.37c}$$

Monte Carlo simulations were performed to validate the approximated model.



(a) Mean vs. time.



(b) Standard deviation vs. time.

Fig. 10.: (a) Mean and (b) standard deviation of the true system (solid) and the approximated system (dashed) obtained from Monte Carlo simulations.

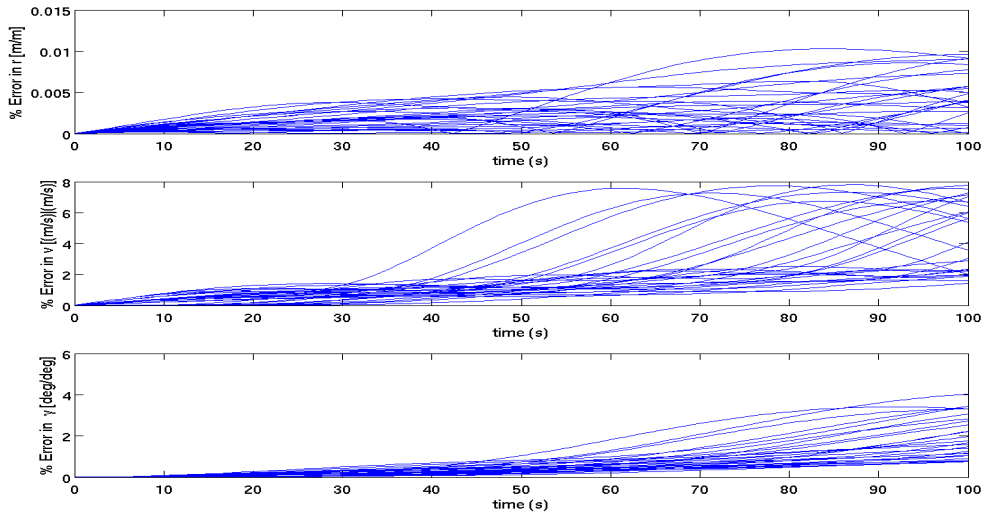


Fig. 11.: Percentage error in states of the approximated system from true system.

A Gaussian distribution for initial states was assumed, with mean and standard deviation as given in Eqn. (3.36a) & Eqn. (3.36b). Figure 10 shows the trajectories of mean and the standard deviations for systems in Eqn. (3.34a) through Eqn. (3.34c) and Eqn. (3.37a) through Eqn. (3.37c). It can be seen that they are almost identical for both the systems, although some variation is observed for longer propagation time. Figure 11 shows the percentage error in approximation of the systems' states. From this analysis it can be concluded that the system in Eqn. (3.37a) through Eqn. (3.37c) is a good approximation of the actual system in Eqn. (3.34a) through Eqn. (3.34c), for estimation purposes. Clearly a better approximation could be performed that reduces the growth of error over long time integration. However, with this model, good results were obtained, for the estimation algorithm. The approximated model is used to propagate the uncertainty. The "measurements" were obtained by simulating the actual nonlinear system.

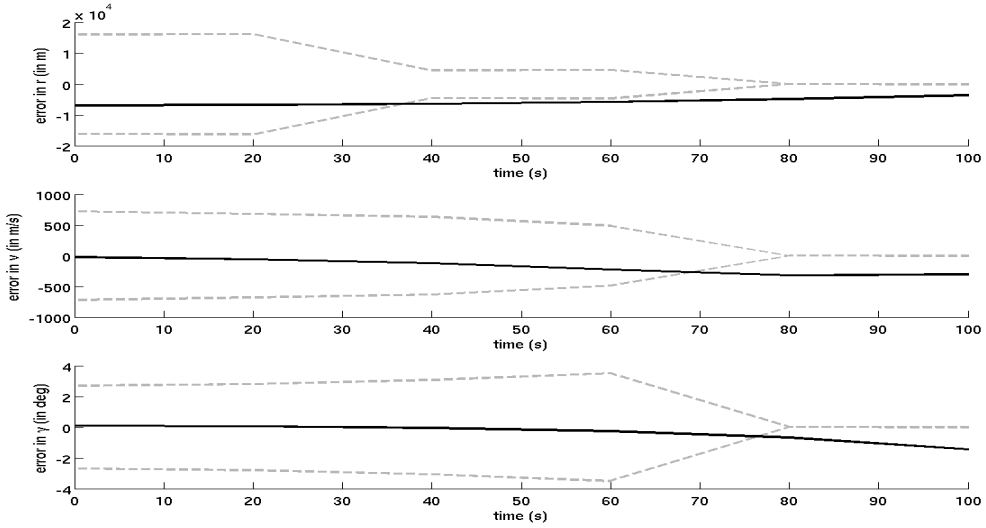
To construct the deterministic dynamical system via polynomial chaos as defined in Eqn. (3.7), inner products of up to six basis functions were required. Detailed derivation of the augmented dynamics has been omitted from this dissertation and can be found in Prabhakar *et al.* [102].

Up to fourth order moments were used, to compute the prior probability density function from GMM. The prior moments, M^{i-} for $i = 1, \ldots, 4$ were computed from the gPC coefficients X_{PC} and the inner products $\langle \phi_i \phi_j \ldots \rangle$ as in Eqn. (3.19a) through Eqn. (3.19d).

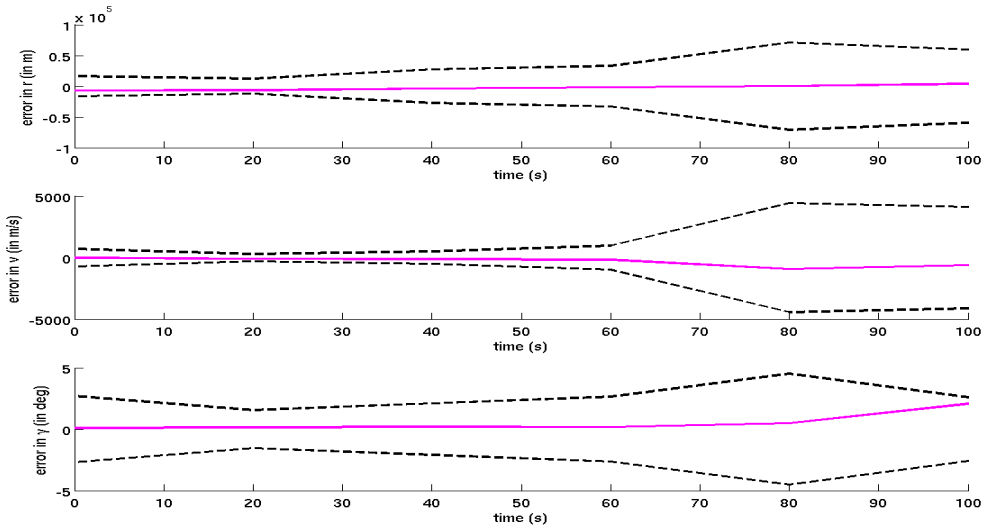
The performance of the proposed estimator was then compared with EKF and UKF based estimators. For the current analysis, the covariance of the posterior PDF was minimized to get the state estimate (MCE criterion), which is the mean of the posterior density function. The initial states of the true system were assumed to be $x(t = 0) = [R_m + 61 \text{ km } 2.42 \text{ km/s } -8.91^o]^T$. Therefore the initial error in state estimation was $[-7 \text{ km } -0.02 \text{ km/s } 0.09^o]^T$. Figure 12 and Fig. 13, show the plots for error in state estimation where the measurement update interval is 20s. It can be observed that the errors for gPC based estimator goes to zero rapidly and are within the $\pm 3\sigma$ limits. The EKF based estimator performs poorly and the errors don't converge to zero. Also the EKF based estimator is inconsistent as the errors escape outside $\pm 3\sigma$ limits. In case of UKF, the errors for γ and v starts to increase after a while and the $\pm 3\sigma$ limits are divergent.

In Fig. 14 and Fig. 15 the update interval was increased to 40s. It can be seen that, the errors and the $\pm 3\sigma$ bounds for the gPC based estimator converges. The errors for the EKF based estimator for this case diverges. Although the $\pm 3\sigma$ limits of the UKF based estimator converge, but convergence is slower than gPC based estimator. The estimation errors for UKF is divergent.

The gPC based estimator works well even for larger errors in initial state es-



(a) EKF based estimator.



(b) UKF based estimator.

Fig. 12.: Performance of (a) EKF and (b) UKF based estimator with true initial states as $[R_m + 61 \text{ km } 2.42 \text{ km/s } -8.91^o]^T$ and update interval of 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

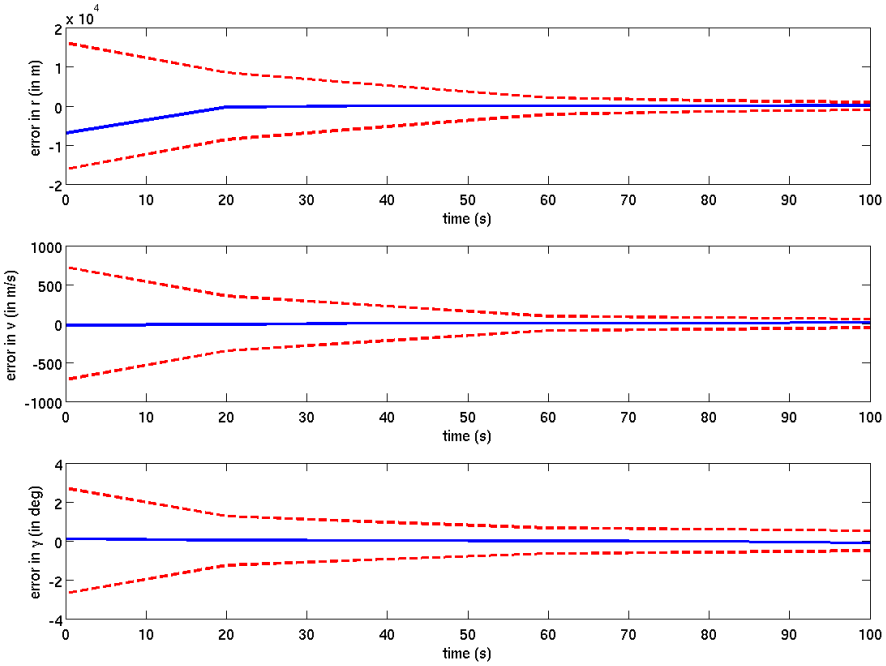
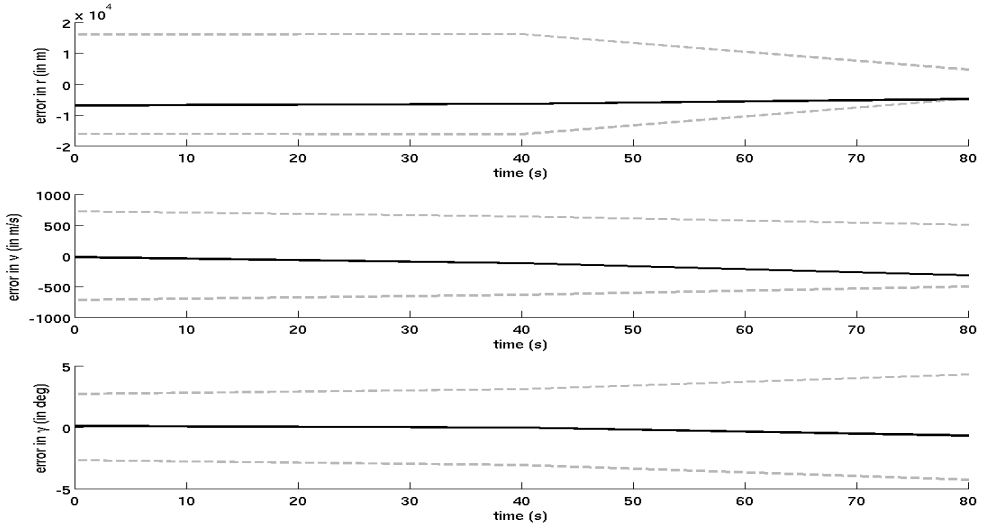
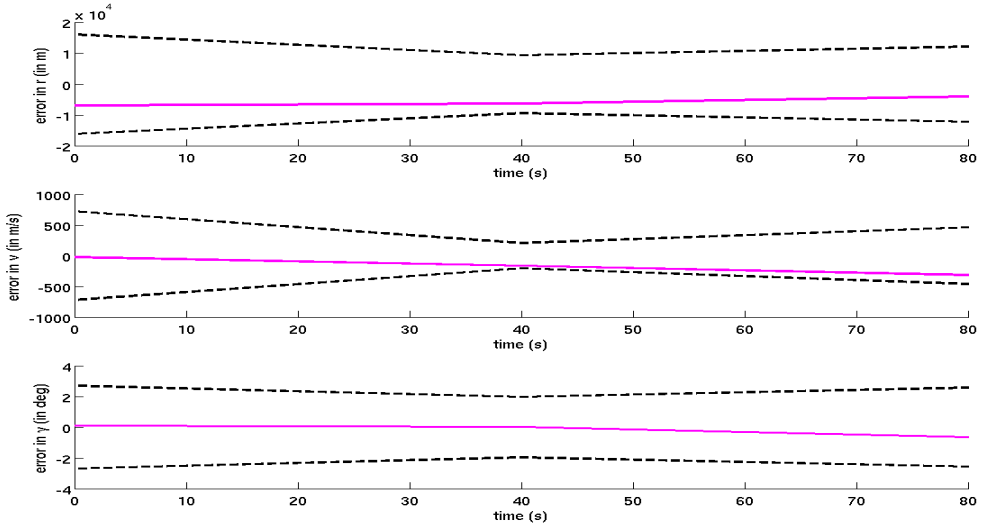


Fig. 13.: Performance of the gPC based estimator with true initial states as $[R_m +$ 61 km 2.42 km/s $-8.91^o]^T$ and update interval of 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.



(a) EKF based estimator.



(b) UKF based estimator.

Fig. 14.: Performance of (a) EKF and (b) UKF based estimator with true initial states as $[R_m + 61 \text{ km } 2.42 \text{ km/s } -8.91^o]^T$ and update interval of 40s.The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

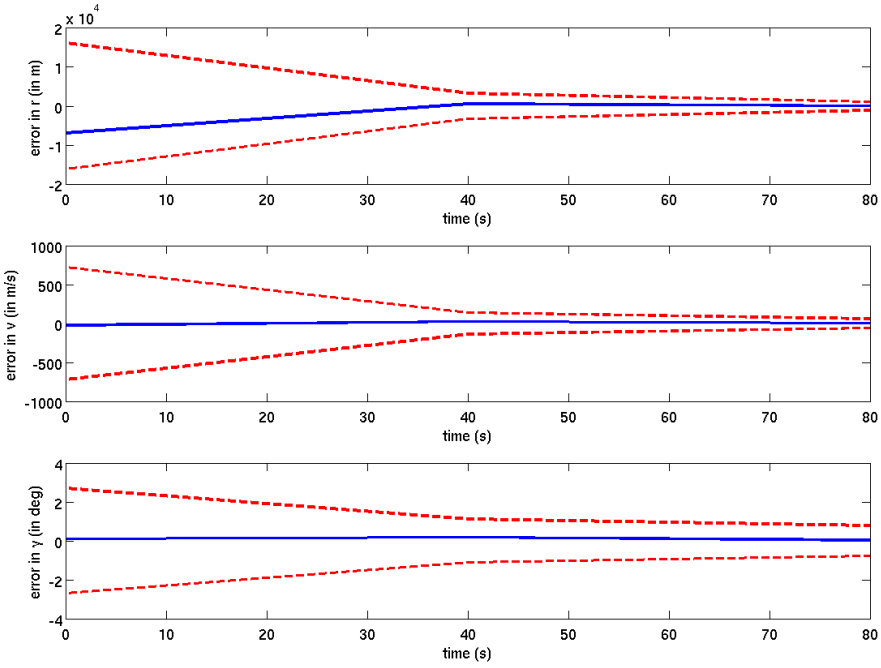
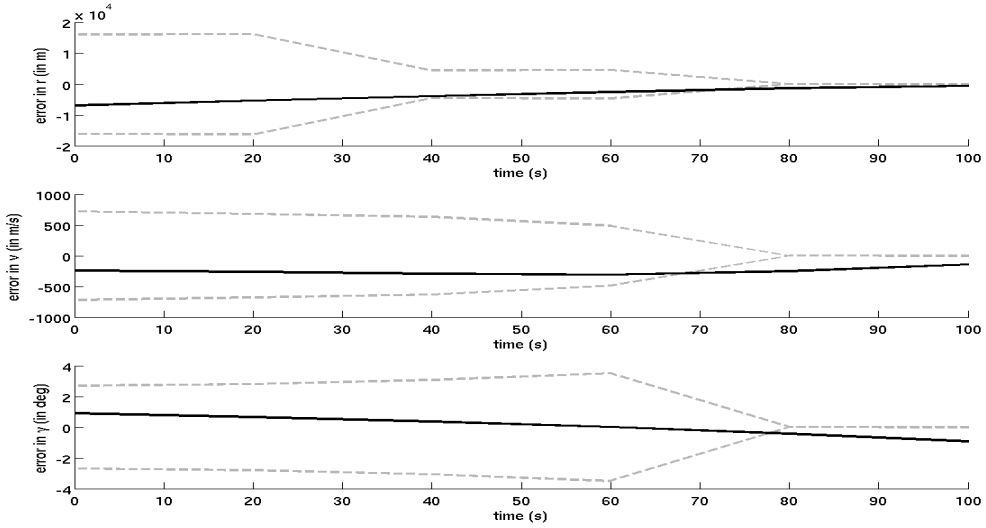


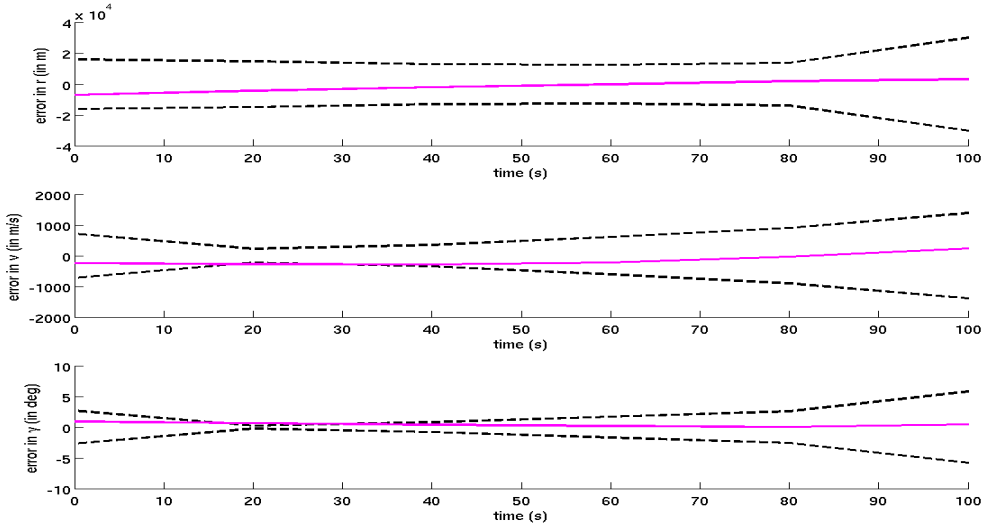
Fig. 15.: Performance of the gPC based estimator with true initial states as $[R_m +$ 61 km 2.42 km/s $-8.91^o]^T$ and update interval of 40s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

timation. For plots in Fig. 16 and Fig. 17, the true initial states were $[R_m +$ 61 km 2.64 km/s $-8.1^o]^T$, with v and γ having $10\%(=\sigma)$ errors in initial estimate. It can be observed that gPC based estimator achieves convergence of error and remains within $\pm 3\sigma$ bounds. For EKF, the errors diverge and escape outside the $\pm 3\sigma$ limits. Hence the EKF based estimator here is inconsistent. The errors for the UKF based estimator converge but briefly remain outside the $\pm 3\sigma$ bounds. For this example, the $\pm 3\sigma$ limits for UKF are observed to be divergent. This highlights the advantage of using polynomial chaos for propagation in Bayesian estimation framework over traditional filters based on Gaussian theory.

Figure 18 compares the performance of the three criterion functions used to get the state estimate. It is observed that $\pm 3\sigma$ bounds of MEE criterion converges faster than MCE and MLE criteria, although the bounds for MLE based estimate diverges towards the end. The error convergence for MEE is slower MCE and MLE, but errors are within $\pm 3\sigma$ limits for all the criteria. When compared to EKF and UKF, the gPC based estimators outperform both of them when convergence of errors and $\pm 3\sigma$ limits are considered.



(a) EKF based estimator.



(b) UKF based estimator.

Fig. 16.: Performance of (a) EKF and (b) UKF based estimator with true initial states as $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$ and update interval of 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

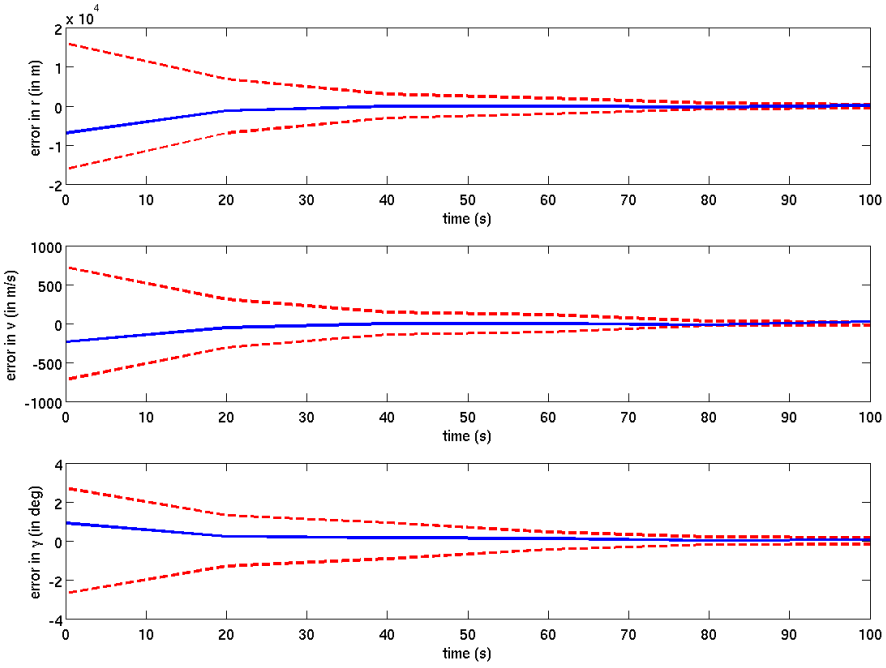
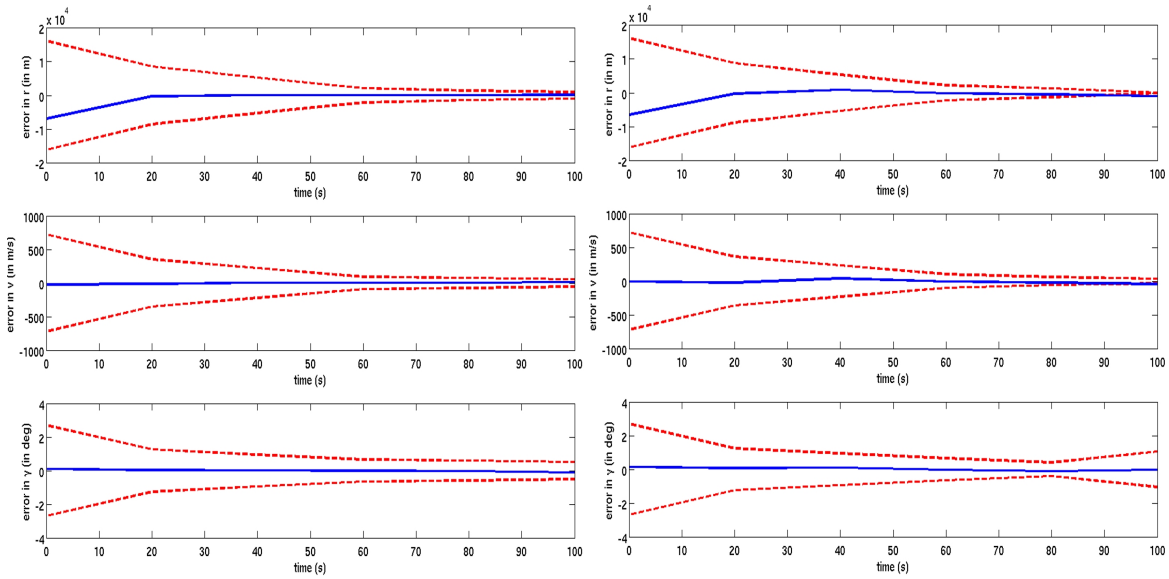
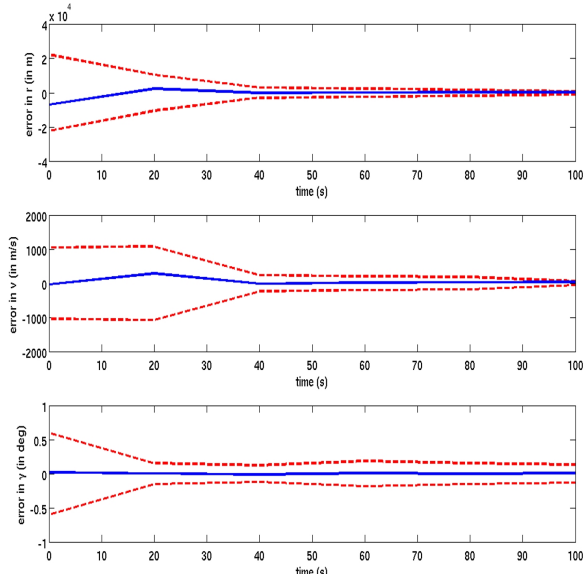


Fig. 17.: Performance of the gPC based estimator with true initial states as $[R_m +$ 61 km 2.64 km/s $-8.1^o]^T$ and update interval of 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.



(a) MCE criterion.

(b) MLE criterion.



(c) MEE criterion.

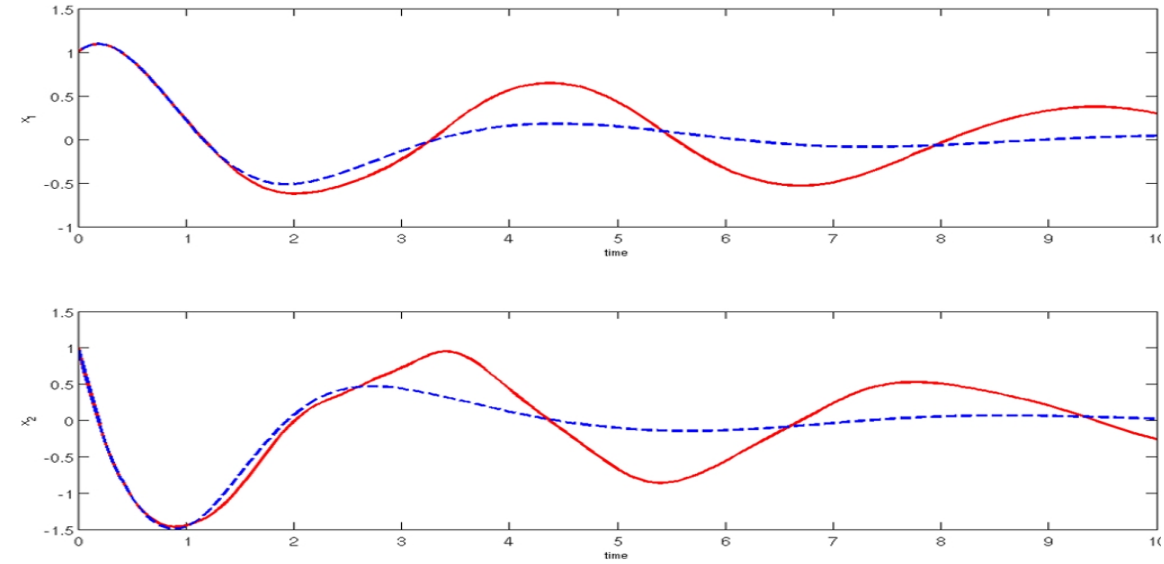
Fig. 18.: Performance of (a) MCE, (b) MLE and (c) MEE criterion for state estimates. The true initial states are $[R_m + 61 \text{ km } 2.42 \text{ km/s } -8.91^o]^T$ and update interval is 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

G. Limitations of Polynomial Chaos

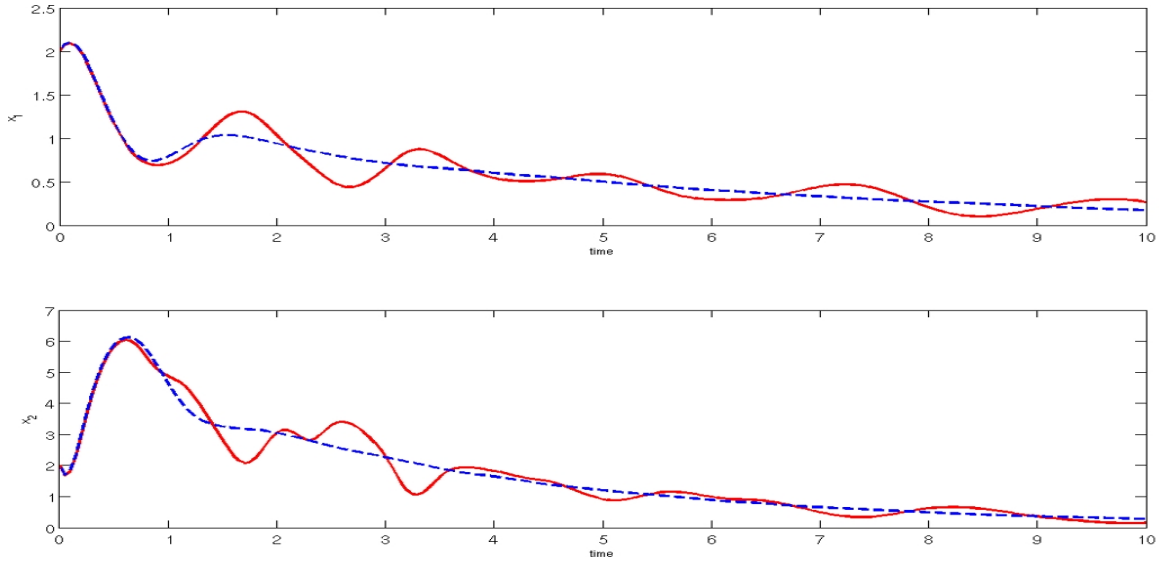
Polynomial chaos is a good methodology to propagate uncertainty specially in systems where nonlinearities are in the form of polynomials. But there are certain limitations, when we use gPC to propagate uncertainty in systems which have non-polynomial nonlinearities. This is well documented and can be found in the paper by Debusschere *et al.* [87]. Moreover polynomial chaos gives us inaccurate results when the statistics is to predicted for longer durations.

For example Fig. 19 shows first and second moment plots for the system given by Eqn. (3.10) & Eqn. (3.11) with Gaussian initial state uncertainty have mean and covariance, $\mu = [1, 1]$ and $\sigma = \text{diag}(1, 1)$ respectively. It can be seen that after a while the PC trajectory diverges from the MC solution, thus proving that PC incurs error when long term statistics are to be predicted.

Uncertainty propagation using PC is a method, which is based on parameterizing the random variable, and then approximating the density using a density kernel. However, a method which propagates density directly without requiring any kind of parameterization is more desired. Intuitively this nonparametric method should perform better than PC as it is approximation free. In the next chapter, we present a methodology to propagate uncertainty in a nonparametric sense.



(a) First moment.



(b) Second moment.

Fig. 19.: Long term statistics, up to 10 seconds predicted by Monte Carlo (dashed line) and polynomial chaos (solid line) for the system in Eqn. (3.10) & Eqn. (3.11). The x-axis represents time and y-axes are the states.

CHAPTER IV

THE FROBENIUS-PERRON OPERATOR*

In this chapter, we propose a nonlinear estimation technique based on Frobenius-Perron (FP) operator theory. The Frobenius-Perron operator, also known by the name of transfer operator or Stochastic Liouville equation (SLE) is a linear operator that dictates the flow of probability densities in a dynamical system. We introduce FP operator, then develop theories related to how densities evolve in a dynamical system with uncertainties present in it. Finally we describe the estimation algorithm, and apply it to hypersonic reentry problem, while comparing its performance with particle filters.

A. Analysis of Densities in State Space Using Frobenius-Perron Operator

In this section we introduce the concept of the FP operator and demonstrate, how it can be used to study evolution of densities of state trajectories. The presentation of the background on FP operator is deliberately kept informal as the associated theory is well developed. The interested reader is directed to [26] for a rigorous treatment of the material presented in this section.

1. Discrete-Time Systems

To establish an intuitive view of the evolution of densities, let us consider one dimensional discrete-time system defined by the transformation $x^{k+1} = S(x^k)$, where $S : \mathcal{A} \mapsto \mathcal{A}$, and $\mathcal{A} \subset \mathbb{R}$. Let us pick a large number of N initial states $x_1^0, \cdots, x_N^0$

*Reprinted from "Hypersonic State Estimation Using Frobenius-Perron Operator" by P. Dutta, R. Bhattacharya, 2011. *AIAA Journal of Guidance Control and Dynamics*, vol. 34, no. 2, pp. 325–344, Copyright [2011] by P. Dutta & R. Bhattacharya

and apply the map $S(x)$ to obtain new states $x_1^1 = S(x_1^0), \cdots, x_N^1 = S(x_N^0)$. Let us define an indicator function for $\mathcal{A}$ as

$$\mathbf{1}_{\mathcal{A}}(x) = \begin{cases} 1 \text{ if } x \in \mathcal{A}, \\ 0 \text{ if } x \notin \mathcal{A}. \end{cases}$$

Also any function $f_0(x)$ is a density function. for initial states $x_1^0, \cdots, x_N^0$, if for every interval $\mathcal{D}_0 \subset \mathcal{A}$ and $\mu(\mathcal{D}_0) \neq 0$, the following is true,

$$\int_{\mathcal{D}_0} f_0(u)du = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{\mathcal{D}_0}(x_i^0), \tag{4.1}$$

Similarly the density function $f_1(x)$ for states $x_1^1, \cdots, x_N^1$ satisfies, for $\mathcal{D}_1 \subset \mathcal{A}$,

$$\int_{\mathcal{D}_1} f_1(u)du = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{\mathcal{D}_1}(x_i^1), \tag{4.2}$$

The FP operator relates f_1 and f_0. To define the FP operator, we first introduce the notion of *counterimage* of any interval $\mathcal{D} \subset \mathcal{A}$ under transformation operator S. This is the set of all points that will be in $\mathcal{D}$ after one application of S, i.e.,

$$S^{-1}(\mathcal{D}) = \{x : S(x) \in \mathcal{D}\}.$$

We note that $\forall \mathcal{D}_1 \subset \mathcal{A}$, we have

$$x_j^1 \in \mathcal{D}_1 \iff x_j^0 \in S^{-1}(\mathcal{D}_1).$$

Hence $\mathcal{D}_0 = S^{-1}(\mathcal{D}_1)$ as $x_j^0 \in \mathcal{D}_0$. We thus have an important relation, $\forall \mathcal{D} \subset \mathcal{A}$.

$$\mathbf{1}_{\mathcal{D}}(S(x)) = \mathbf{1}_{S^{-1}(\mathcal{D})}(x). \tag{4.3}$$

With Eqn. (4.3) we can write Eqn. (4.2) as

$$\begin{aligned}\int_{\mathcal{D}_1} f_1(u)du &= \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} 1_{S^{-1}(\mathcal{D}_1)}(x_i^0) \\ &= \int_{S^{-1}(\mathcal{D}_1)} f_0(u)du.\end{aligned}$$

The above relation tells us how density of initial states f_0 will be transformed to new density f_1 under the transformation S.

Let $\mathcal{D}_1 = [a, x]$ then,

$$\int_a^x f_1(u)du = \int_{S^{-1}[a,x]} f_0(u)du,$$

and differentiating the above relation with respect to x, we have

$$f_1(x) = \frac{d}{dx}\int_{S^{-1}[a,x]} f_0(u)du$$

Clearly, f_1 depends on f_0, and is usually indicated by writing $f_1 = Pf_0$. In general, the deformation of the density function $f(x)$ defined over $[a, x]$ under transformation S is given by

$$Pf(x) = \frac{d}{dx}\int_{S^{-1}[a,x]} f(u)du \tag{4.4}$$

If the transformation S is invertible and differentiable in $\mathcal{A}$, then it must be monotone in $\mathcal{A}$. In such case Eqn. (4.4) can be written as

$$\begin{aligned}Pf(x) &= \frac{d}{dx}\int_{S^{-1}[a,x]} f(u)du \\ &= f(S^{-1}(x))\left|\frac{dS^{-1}(x)}{dx}\right|,\end{aligned} \tag{4.5}$$

where absolute value takes into account cases where S^{-1} is increasing or decreasing. Equation (4.5) can be generalized for $x \in \mathbb{R}^n$ by noting that $|dS^{-1}(x)/dx|$ is the determinant of the Jacobian of $dS^{-1}(x)/dx$. Time evolution of $f(x)$ can be determined by recursively applying Eqn. (4.4).

2. Continuous-Time Systems

In contrast to discrete time systems, we have *uncountably many* mappings from initial to final states, in case of continuous time systems. We will start by defining some preliminary concepts, on which the development of FP operator theory is based.

Let $\mathcal{X}$ be a Hausdorff space and $\mathcal{A}$ be the $\sigma-$algebra generated by Borel sets of $\mathcal{X}$.

Definition 1 *A dynamical system $\{S_t\}_{t\in\mathbb{R}}$ on $\mathcal{X}$ is the family of transformations $S_t : \mathcal{X} \to \mathcal{X}, t \in \mathbb{R}$, satisfying*

- $S_0(x) = x \quad \forall x \in \mathcal{X}$
- $S_t(S_{t'}(x)) = S_{t+t'}(x) \quad \forall x \in \mathcal{X}$ *and* $t, t' \in \mathbb{R}$, *and*
- *The mapping $(t, x) \to S_t(x)$ from $\mathbb{R} \times \mathcal{X} \to \mathcal{X}$ is continuous.*

It is also clear from the properties of $\{S_t\}_{t\in\mathbb{R}}$ that $S_t(S_{-t}(x)) = x$ and $S_{-t}(S_t(x)) = x$, $\forall t \in \mathbb{R}$. Thus, $\forall t_0 \in \mathbb{R}$, any transformation S_{t_0} of the dynamical system $\{S_t\}_{t\in\mathbb{R}}$, is invertible.

Remark 1 *A semidynamical system is same as a dynamical system except for $t \in \mathbb{R}^+$.*

Let us consider a semidynamical system of the form $\{S_t\}_{t\geq 0}$. Assume a measure $\mu(\cdot)$ on $\mathcal{X}$ is given and all transformation S_t of $\{S_t\}_{t\geq 0}$ are nonsingular, i.e.,

$$\mathcal{F} = \{A \in \mathcal{A} : \mu(S_t^{-1}(A)) = 0\} \text{ then } \mu(A) = 0 \quad \forall A \in \mathcal{F}.$$

Definition 2 *The Frobenius-Perron operator $P_t : L^1(\mathcal{X}) \to L^1(\mathcal{X})$, corresponding to the transformation S_t is uniquely defined by*

$$\int_A P_t f(x)\mu(dx) = \int_{S^{-1}(A)} f(x)\mu(dx),$$

where $f : \mathcal{X} \to \mathbb{R}$ is any arbitrary function such that $\int_{\mathcal{X}} |f(x)|\mu(dx) < \infty$.

We can show [26] that $\{P_t\}_{t\geq 0}$ is continuous. To show this, note that since S_t is invertible, $S_t^{-1} = S_{-t}$. Let J_{-t} be the determinant of the Jacobian of the transformation S_{-t}. Then, using Eqn. (4.5)

$$P_t f(x) = f(S_{-t}(x))J_{-t}(x).$$

Thus, for every continuous f, with compact support,

$$\lim_{t\to t_0} f(S_{-t}(x))J_{-t}(x) = f(S_{-t_0}(x))J_{-t_0}(x),$$

uniformly with respect to x. This implies that,

$$\lim_{t\to t_0} ||P_t f - P_{t_0} f|| = \lim_{t\to t_0} \int_X |P_t f(x) - P_{t_0} f(x)|dx = 0,$$

Since the integrals are computed over the a bounded set $X \subset \mathcal{A}$. Now, because continuous functions with compact support form a dense subset in $\mathcal{L}^1$, this implies that $\{P_t\}_{t\geq 0}$ is continuous.

From a computational point of view, this property is important as we will be using space of continuous functions to approximately solve for $\{P_t\}_{t\geq 0}$ in $\mathbb{R} \times \mathcal{X}$.

For a dynamical system of the form

$$\dot{x} = F(x);\ x \in \mathcal{X} \in \mathcal{B}(\mathbb{R}^n),\ F : \mathcal{X} \to \mathcal{X}, \tag{4.6}$$

where $\mathcal{B}(\mathbb{R}^n)$ refers to Borel sets of $\mathbb{R}^n$ and $F(x) = [F_1(x), \cdots, F_n(x)]^T$. It can

be shown that, the definition of the FP operator reduces to the following partial differential equation (pg. 213, [26]).

$$\frac{\partial u}{\partial t} + \sum_{i=1}^{n} \frac{\partial (uF_i(x))}{\partial x_i} = 0, \tag{4.7}$$

where $u(t, x) = P_t f(x)$.

3. Properties of Frobenius-Perron Operator

Both continuous and discrete time Frobenius-Perron operators have the following properties, shown here using the notation for continuous time FP operator,

1. $P_t(\lambda_1 f_1 + \lambda_2 f_2) = \lambda_1 P_t f_1 + \lambda_2 P_t f_2,\ \forall f_1, f_2 \in \mathcal{L}^1,\ \lambda_1, \lambda_2 \in \mathbb{R}$.

2. $P_t f \geq 0$, if $f \geq 0$.

3. $\int_{\mathcal{X}} P_t f(x)\mu(dx) = \int_{\mathcal{X}} f(x)\mu(dx),\ \forall f \in \mathcal{L}^1$.

   Thus for a fixed t, P_t is a Markov operator.

B. Method of Characteristics for Solving First Order Linear Partial Differential Equations

Equation (4.7) is a first order linear partial differential equation that can be solved using method of characteristics. Equation (4.7) can be written as

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x_1} F_1(x) + \cdots \frac{\partial u}{\partial x_n} F_n(x) + u \sum_{i=1}^{n} \frac{\partial F_i(x)}{\partial x_i} = 0.$$

Defining $G(x, u) := -u \sum_{i=1}^{n} \frac{\partial F_i(x)}{\partial x_i}$, we get

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x_1} F_1(x) + \cdots \frac{\partial u}{\partial x_n} F_n(x) = G(x, u), \tag{4.8}$$

which is in the standard form. Assuming $G(x,u) \neq 0$, Eqn. (4.8) can be solved by solving $(n+1)$ coupled ordinary differential equations given by

$$\frac{dx_1}{dt} = F_1(x), \cdots, \frac{dx_n}{dt} = F_n(x), \tag{4.9a}$$

$$\frac{du}{dt} = G(x,u). \tag{4.9b}$$

Substituting the value of $G(x,u)$ in Eqn. (4.9b), we get

$$\frac{du}{dt} = -u \sum_{i=1}^{n} \frac{\partial F_i(x)}{\partial x_i} \tag{4.10}$$

where $\sum_{i=1}^{n} \frac{\partial F_i(x)}{\partial x_i}$ is the trace of the Jacobian of $F(x)$.

Equation (4.9a) & Eqn. (4.9b) trace out a trajectory in the $(n+1)$ dimensional space spanned by $(x_1, \cdots, x_n, u)$. To make the solution unique, the value of $u(t,x)$ has to be specified at a given point $x(t_0)$ at time t_0. The evolution of $u(t,x)$ over $\mathbb{R} \times \mathcal{X}$ can be determined by specifying $u(t,x)$ over several points in $\mathcal{X}$ at time t_0. The trajectories $x(t)$ are called *characteristics* and Eqn. (4.9a) & Eqn. (4.9b) determine evolution of u along $x(t)$ [103].

C. Uncertainty Propagation in Dynamical Systems Using FP Operator

Given a dynamical system of the form,

$$\dot{x}(t) = g(x(t), \Delta) \tag{4.11}$$

where $x : \mathbb{R}^+ \to \mathbb{R}^n$ are states and $g : \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}^n$ is a real valued function. Let $\Delta \in \mathbb{R}^q$ be the vector of initial and parametric uncertainty. Without loss of generality, we assume that Δ represents the set of initial state uncertainty, hence $\Delta := x(t=0) \in \mathbb{R}^n$, and let $\Delta \sim p(\Delta) = p(x(0))$. Using the method of characteristics to get the equation for evolution of densities as presented in the previous section in

Eqn. (4.10), we augment the dynamical system in Eqn. (4.11). The final dynamical system is given by,

$$\dot{x}(t) = g(x(t)) \tag{4.12a}$$

$$\dot{p}(t) = -\operatorname{div}\left[g(x(t))\right] p(t), \tag{4.12b}$$

where $p : \mathbb{R}^+ \to \mathbb{R}$ denotes the joint density of the states $x(t)$ at a given time t. The augmented equation has states $[x(t), p(t)]$, which are to be solved with initial conditions $[x(t=0), p(t=0)]$. Equation (4.12a) & Eqn. (4.12) dictate the evolution of probability density along the trajectories of a dynamical system with uncertainties.

For example, consider uncertainty propagation due to initial condition uncertainty for the Duffing oscillator model. As shown in Fig. 20, consider the case where there is initial condition has the mean $(1, 1)$ and standard deviation $\mathrm{diag}(1, 1)$. The distribution is considered to be Gaussian. The initial distribution is sampled using 500 points. The system considered is the same one as used in Eqn. (3.10) & Eqn. (3.11) . The corresponding differential equation for uncertainty propagation is given by,

$$\dot{x}_1 = x_2, \tag{4.13a}$$

$$\dot{x}_2 = -x_1 - x_1^3 - \frac{1}{4}x_2, \tag{4.13b}$$

$$\dot{p} = -p(0 + -\frac{1}{4}) = \frac{1}{4}p. \tag{4.13c}$$

The initial conditions for (x_1, x_2) are given by the locations of the samples. The initial conditions for p are given by the values of the Gaussian distribution function at those points. With these initial conditions the above equations are integrated to obtain $(x_1(t), x_2(t), p(t))$. Figure 20 shows the initial values of $(x_1(t), x_2(t), p(t))$ and their values at time $t =$ 1s,2s and 3s. The magnitude of p is color coded, with

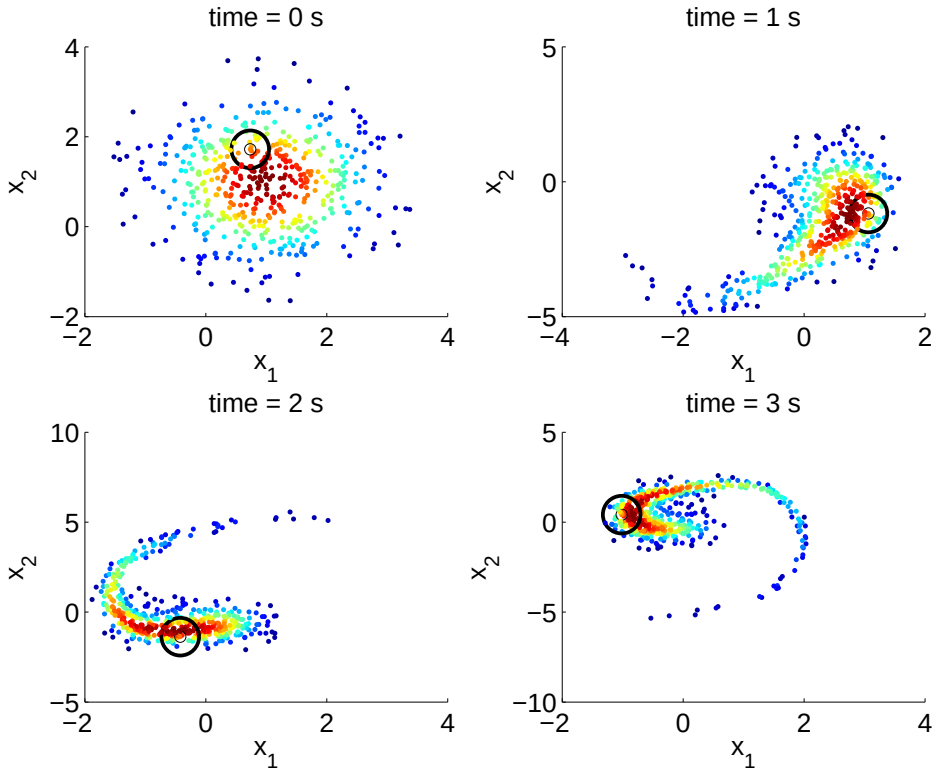
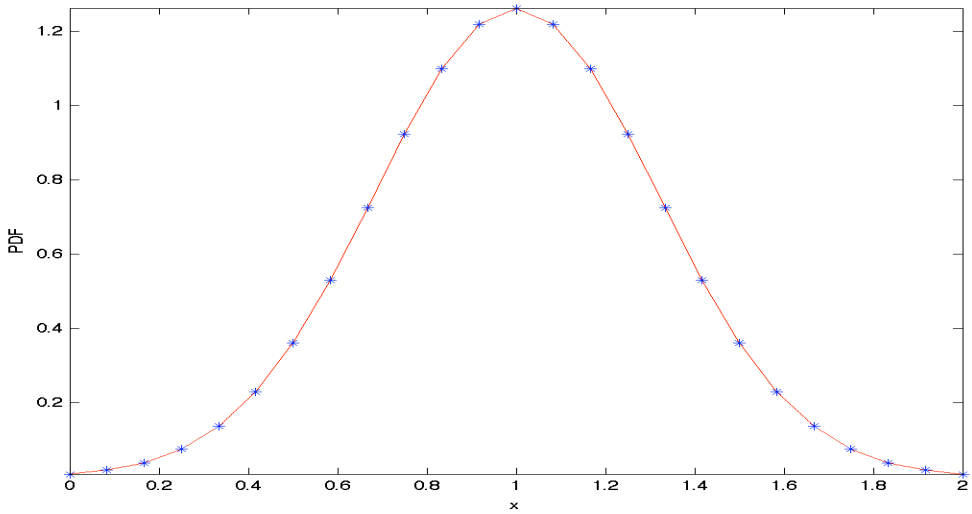


Fig. 20.: Method of characteristics. Number of samples = 500. System is Duffing oscillator. The black circle shows the location and the PDF value (color coded) of an arbitrary sample point at different time instances.

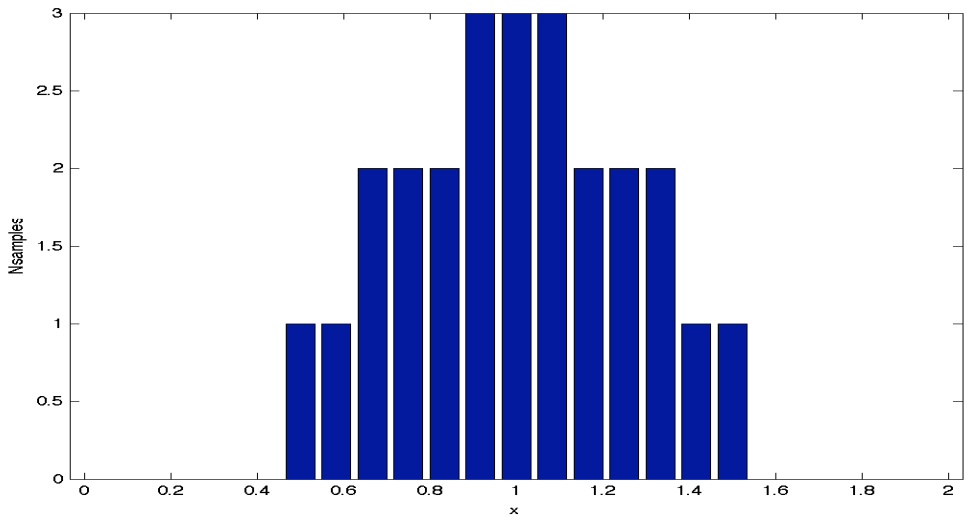
red corresponding to large and blue corresponding to small. It can be observed, this method determines the value of $p(t)$ at the sample locations, (x_1, x_2) exactly. However, for intermediate points, the value of $p(t)$ is not known. Several methods like scattered data interpolation [104] may be used to find the density at intermediate points, but such an analysis has not been performed in this dissertation. Also, note that value of p determined at (x_1, x_2), at some time t, is independent of the number of samples. Although, increasing the number of samples will improve the accuracy of the subsequent interpolation. However, in case of Monte Carlo, since the value of p is approximated using histograms, the accuracy depends on the number of samples.

1. Accuracy in Prediction of Uncertainty

Since this method requires selection of points in the state space, one may argue that this approach is similar to Monte Carlo. The main difference is that in FP operator the value of the density function is determined along the trajectory and at final time the value of the density function is known at certain discrete locations. These locations are values of the state vector at that final time. The value of the density function over the domain can be then determined using interpolation. In Monte Carlo, same idea is used except only state equations are integrated. The density function at final time is determined using *histogram* techniques, which are sensitive to the number of sample points taken. Thus for comparable accuracy, many more points will have to be taken for Monte Carlo based method than for FP based method. As an illustration, consider Fig. 21. It can be seen that a good approximation of the Gaussian density function is achieved by taking 25 points and linear interpolation in the case of FP operator. Whereas for Monte Carlo, at least 1000 points were needed to get the same level of accuracy. Thus for this one dimensional case, there is a reduction of two orders of magnitude in the number of sample points required for FP based uncertainty propagation. For higher dimensional case, the benefits will be even more significant. For high dimensional problems, discrete points can be generated using Halton sequence [105], and concentrating them in regions with higher importance/features will further reduce the growth of computational complexity. Other methods for generating smart samples from a given distribution include sparse grids [106, 107, 58] and mesh-free adjoint methods [108, 109, 110, 111].

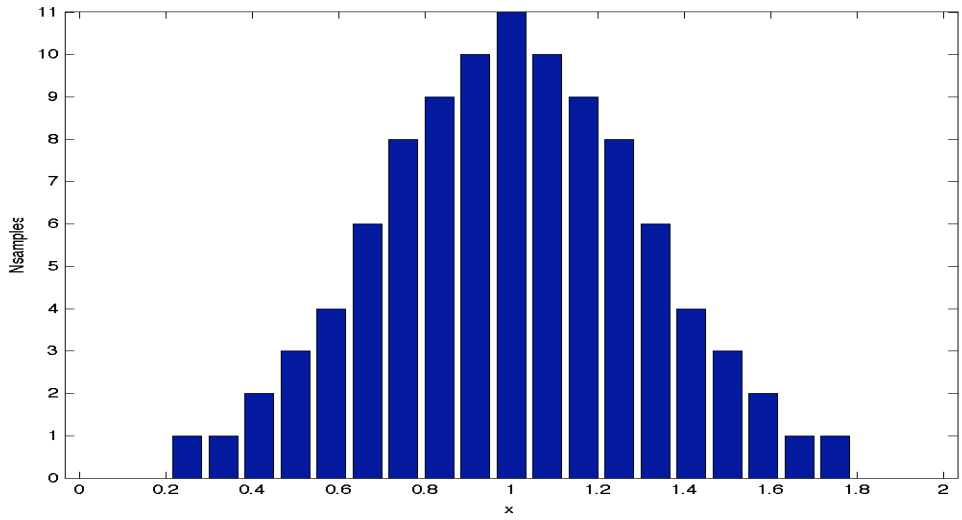


(a) PF: 25 samples.

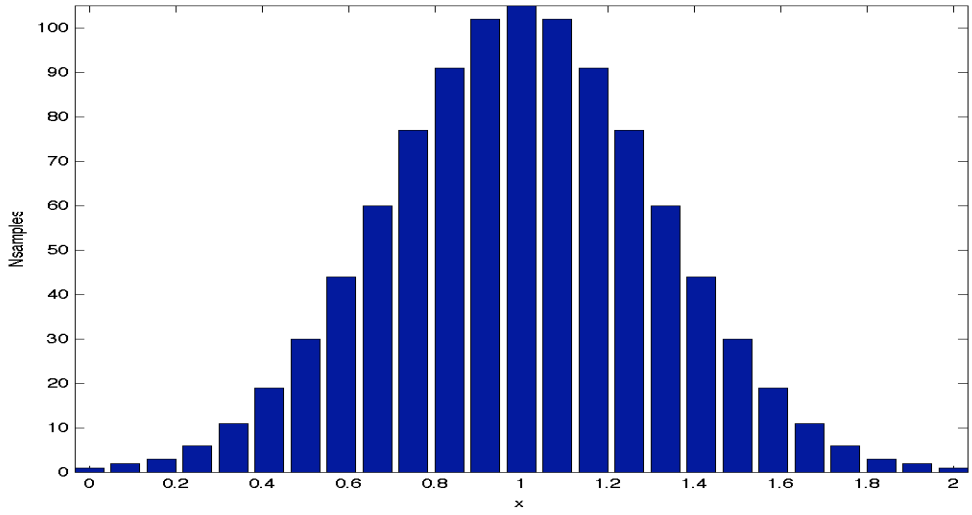


(b) MC: 25 samples.

Fig. 21.: Comparison of FP and Monte Carlo based approximation of density functions.



(c) MC: 100 samples.



(d) MC: 1000 samples.

Fig. 21.: Continued.

D. Nonlinear Estimation Using Frobenius-Perron Operator

The nonlinear estimation setting in this scenario is same as one in Chapter III. We have an uncertain dynamical system being measured by a nonlinear measurement model which are given by,

$$\dot{x} = g(x, \Delta), \tag{4.14a}$$

$$\tilde{y} = h(x) + \nu. \tag{4.14b}$$

where $x \in \mathbb{R}^N$ are the states and $\tilde{y} \in \mathbb{R}^m$ are the measurements.

It is assumed that measurements are available at discrete times $t_0, \cdots, t_k, t_{k+1}, \cdots$. At a given time t_k, let x_k, y_k and $p_k(\cdot)$ be the state, measurement and the probability density function. Let $p_k^-(\cdot)$ and $p_k^+(\cdot)$ denote the prior and the posterior density functions at time t_k. The estimation algorithm is described next, which is similar to that presented by Daum *et al.* [40].

1. Step 1: Initialization Step

To begin, the domain $\mathbf{D}_\Delta$ of the initial random variable x_0 is discretized. From the discretized domain, N particles are chosen at random based on the PDF $p_0(x_0)$ of the random variable x_0. Let the particles be represented by $x_{0,i}$ for $i = 1, 2, \cdots, N$; and $p_0(x_{0,i})$ be the value of $p_0(x)$ at these particles. The following steps are then performed recursively starting from $k = 1$.

2. Step 2: Propagation Step

With initial states at $k - 1^{th}$ step as $[x_{k-1,i}\ p_{k-1}(x_{k-1,i})]^T$ Eqn. (4.9a) & Eqn. (4.9b) are integrated for each particle over the interval $[t_{k-1}, t_k]$ to get $[x_{k,i}, p_k^-(x_{k,i})]^T$. Note that $p_k^-(x_{k,i})$ obtained by integration are the *prior* PDF values for $x_{k,i}$, and hence the

superscript.

3. Step 3: Update Step

First the likelihood function, $p(\tilde{y}_k|x_k = x_{k,i})$, is determined for each particle i, using Gaussian measurement noise and sensor model in Eqn. (4.14b). It is defined as

$$l(\tilde{y}_k|x_k = x_{k,i}) = \frac{1}{\sqrt{(2\pi)^m|R|}} e^{-\frac{1}{2}(\tilde{y}_k - h(x_{k,i}))^T R^{-1}(\tilde{y}_k - h(x_{k,i}))}, \tag{4.15}$$

where $|R|$ is the determinant of the covariance matrix of measurement noise.

The posterior probability density function of the states is constructed next, for each particle i, using classical Bayes rule [78]. It is defined as the density function of the states given current measurement, i.e. $p_k^+(x_{k,i}) := p_k(x_k = x_{k,i}|\tilde{y}_k)$. For a particular particle, $x_{k,i}$ it is given by,

$$p_k^+(x_{k,i}) := p(x_k = x_{k,i}|\tilde{y}_k) = \frac{l(\tilde{y}_k|x_k = x_{k,i})p_k^-(x_k = x_{k,i})}{\sum_{j=1}^{N} l(\tilde{y}_k|x_k = x_{k,j})p_k^-(x_k = x_{k,j})}. \tag{4.16}$$

4. Step 4: Getting the State Estimate

The state estimate for the k^{th} step is then computed depending on the desired computation as given in Bryson and Ho[97]. As in the previous section the commonly used criteria are the same, which are given by,

1. *Maximum-Likelihood Estimate*: maximize the probability that $x_{k,i} = \hat{x}_k$. This results in $\hat{x}_k$ = mode of $p_k^+(x_{k,i})$.

2. *Minimum-Variance Estimate*: here $\hat{x}_k = \arg\min_x \sum_{i=1}^{N} ||x - x_{k,i}||^2 \, p_k^+(x_{k,i}) = \sum_{i=1}^{N} x_{k,i} \, p_k^+(x_{k,i})$. The estimate is the mean of $x_{k,i}$.

3. *Minimum-Error Estimate*: minimize maximum $|x - x_{k,i}|$. This results in $\hat{x} =$ median of $p_k^+(x_{k,i})$.

Note that though the criteria are the same, their definition in this case are different, as we have particles selected from discretized domain of the random variable. Hence we do a *Monte Carlo integration* [112] to get the state estimate.

5. Step 5: Resampling

Degeneracy of particles can be detected by looking at values of $x_{k,i}$ for which $p_k^+(x_{k,i}) < \epsilon$, where $\epsilon \ll 1$ and is pre-specified. Existing methods for resampling can be used to generate new particles from the new distribution $p_k^+(x_{k,i})$. For the $k + 1^{th}$ step, new points and the corresponding posterior density $p_k^+(x_{k,i})$ serve as initial states. Qualitatively, since we have the exact density information for each particle, this method is less sensitive to the issue of degeneracy, as, rather than using histograms, we can directly use the density information. Although, a rigorous analysis of this has not been done in this dissertation.

E. Application of Nonlinear Estimation Using FP Operator

In this section we apply the estimation algorithm using FP operator to estimate states of a nonlinear estimation problem. First we will apply the estimation algorithm to a Duffing oscillator, the will first present results for application to the simplified hypersonic reentry problem given by three state Vinh's equation model. Finally the results of application to the full reentry dynamics will be shown.

1. Application to a Duffing Oscillator System

In this section we consider the Duffing oscillator given in Eqn. (3.10) & Eqn. (3.11). Augmented dynamical system, after application of FP operator, is given by Eqn. (4.13b) & Eqn. (4.13c). We consider initial state uncertainty in x with Gaussian distribution, i.e. $\Delta_{x_0} \sim \mathcal{N}(\mu, \Sigma)$, where,

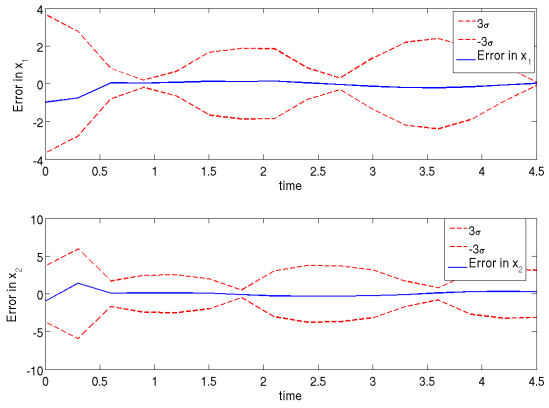
$$\mu = [1, 1]^T \ \Sigma = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}.$$

We use scalar measurement model for the Duffing oscillator which is given by,

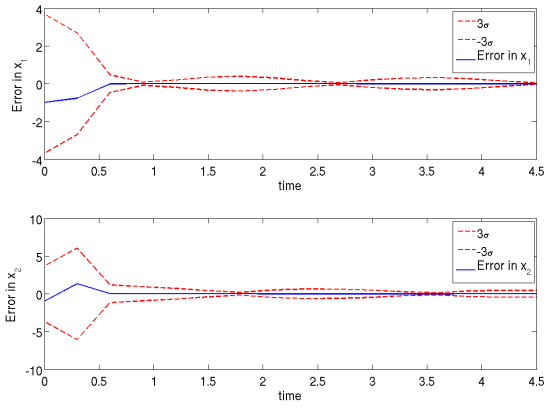
$$\tilde{y} = x^T x + \nu, \tag{4.17}$$

with $\mathbb{E}[\nu] = 0$ and $\mathbb{E}\left[\nu\nu^T\right] = 6 \times 10^{-2}$. Throughout the analysis, the initial states of the **true system** is assumed to be $\bar{x}_0 = [2, 2]^T$.

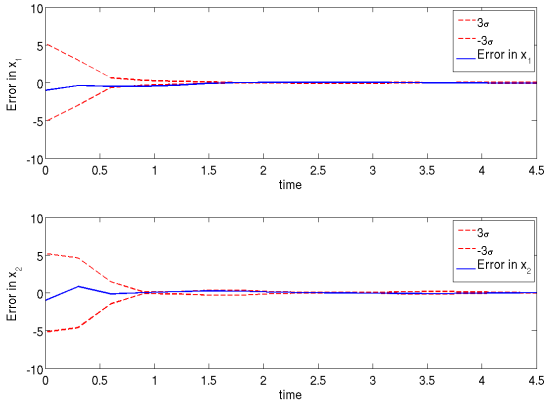
Figure 22 compares the performance of the FP operator based filter with generic particle filters which is described in Chapter II, when the measurement update interval is 0.3s. We plot the estimation error and the $\pm 3\sigma$ bounds here. It can be seen that to achieve similar performance as the FP operator based filter, number of particles needed for the particle filter is more than that of FP operator based filter. In Fig. 23, the update rate is decreased such that each measurement comes after 0.6s interval. It is observed that even in this case, to achieve similar performance the generic particle filter requires more particles. Hence, it can be concluded that for this particular application, FP operator based estimator is computationally better than the generic particle filter. Now we are ready to apply the FP operator based estimation framework to more practical, real-world problems, which has been presented in the subsequent sections.



(a) Particle filter with 100 particles.

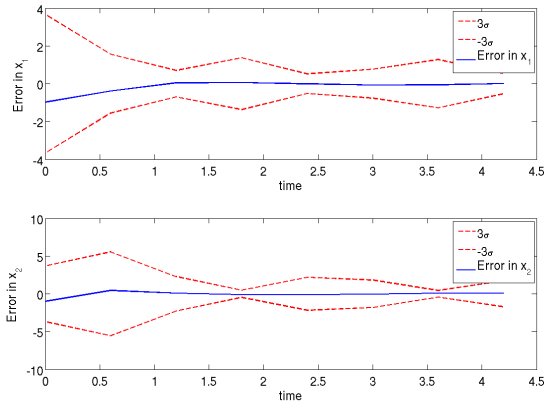


(b) Particle filter with 400 particles.

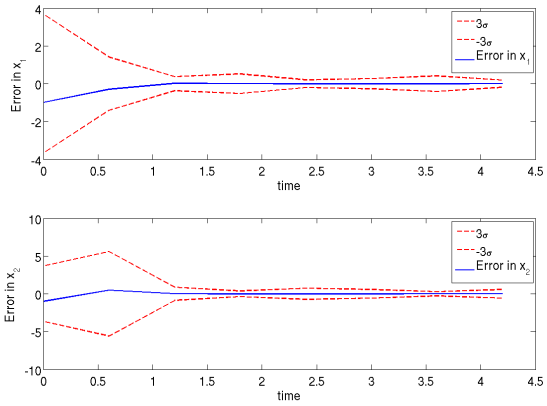


(c) FP operator based filter with 100 particles.

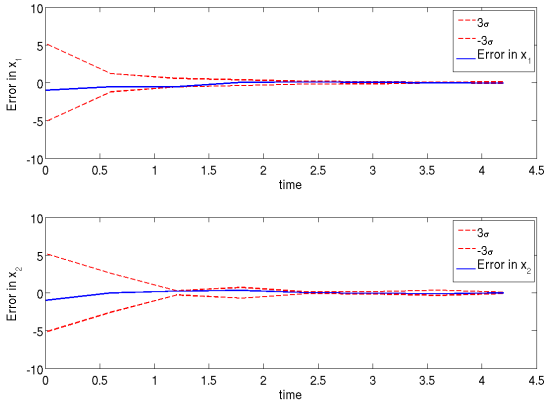
Fig. 22.: Performance of the estimators with measurement update interval as 0.3s.



(a) Particle filter with 100 particles.



(b) Particle filter with 400 particles.



(c) FP operator based filter with 100 particles.

Fig. 23.: Performance of the estimators with measurement update rate as 0.6s.

2. Application to Three State Vinh's Equation

We consider the dynamical system given in Eqn. (3.34a) through Eqn. (3.34c), with states $x = [r, v, \gamma]$, where r is distance from Mars' center, v is the total velocity of the reentry vehicle and γ is the flight path angle. The constants have the same value as given by Table III. The measurement model consist of dynamic pressure $\bar{q}$, heating rate H and the flight path angle γ, whose expressions are given by Eqn. (3.35a) through Eqn. (3.35c). Initial state uncertainty is assumed to be Gaussian with mean μ and standard deviation σ given by Eqn. (3.36a) & Eqn. (3.36b). For consistency in dimensions of the system in Eqn. (3.34a) through Eqn. (3.34c), the base units of mass, length and time are scaled. The scaling factors used are given in Table IV. The measurements are normalized such that they lie in $[-1, 1]$, for consistency in magnitude. The normalization factors used are given in Table V. The measurement noise ν is assumed to have mean and covariance, $\mathbb{E}[\nu] = [0, 0, 0]^T$ and $R = \mathbb{E}\left[\nu\nu^T\right] = 6 \times 10^{-5}\mathcal{I}_3$, respectively, in scaled units.

The performance of the Frobenius-Perron operator based nonlinear estimator is now compared with generic particle filter and bootstrap filter described in Chapter II, when applied to hypersonic reentry dynamics. The initial states of the actual system is taken as $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$, in this case, with v and γ having 10% errors in state estimate. The measurement update rate was taken to be 20 seconds. Figure 24, Fig. 25, and Fig. 26 show the plots for particle filter, the bootstrap filter and the Frobenius-Perron filter respectively. Figure 27 shows the Cramér-Rao bounds for the three filters. From these plots it can be seen that all three filters perform equally well and the errors are within the $\pm 3\sigma$ limits. The number of particles used in these simulations for particle filter, bootstrap filter and the Frobenius-Perron filter are 25000, 20000 and 7000 respectively.

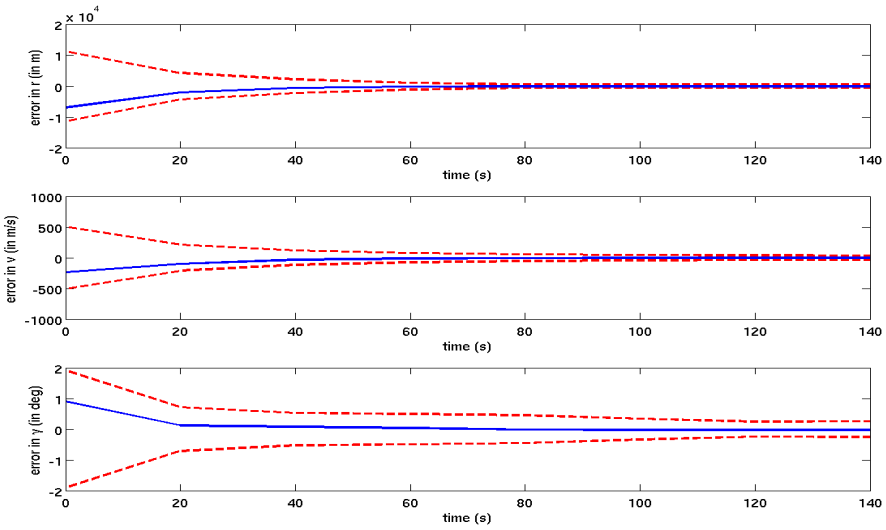


Fig. 24.: Generic particle filter. True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$ and update interval is 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

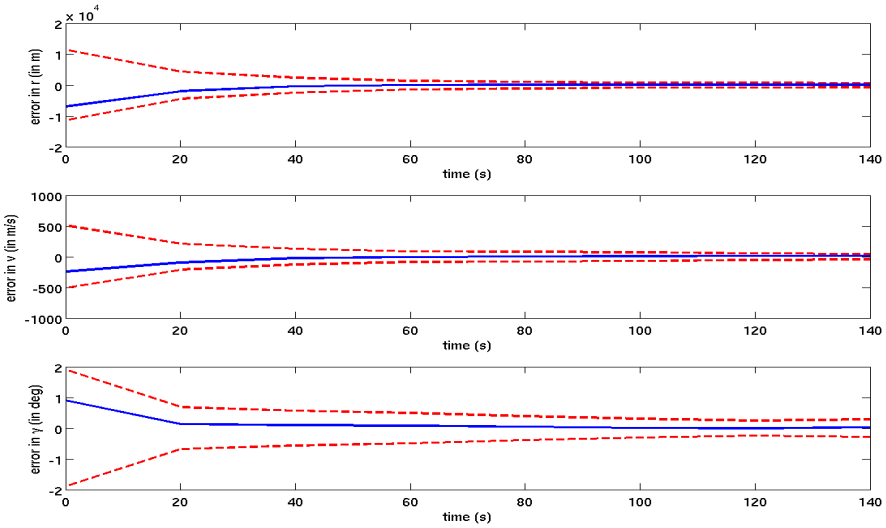


Fig. 25.: Bootstrap filter. True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$ and update interval is 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

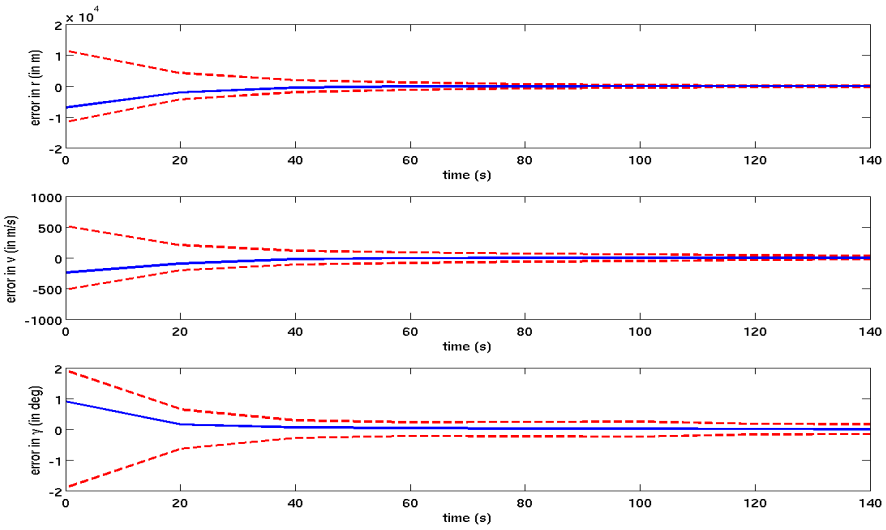


Fig. 26.: Performance of the Frobenius-Perron estimator with true initial states as $[R_m{+}61$ km 2.64 km/s $-8.1^o]^T$ and update interval of 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

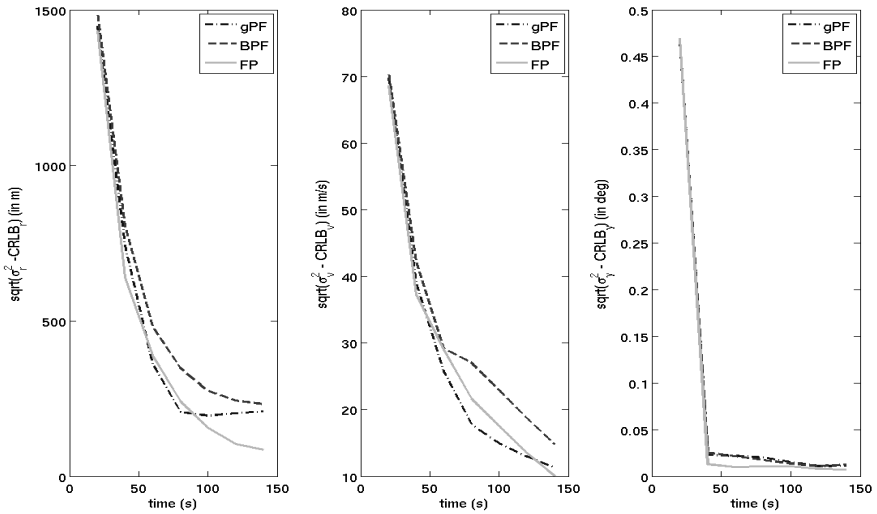


Fig. 27.: Plots for $\sqrt{\sigma^2 - CRLB}$ vs. time. In the legend, 'BPF', 'gPF' and 'FP', represent bootstrap filter, generic particle filter and Frobenius-Perron operator based estimator respectively.

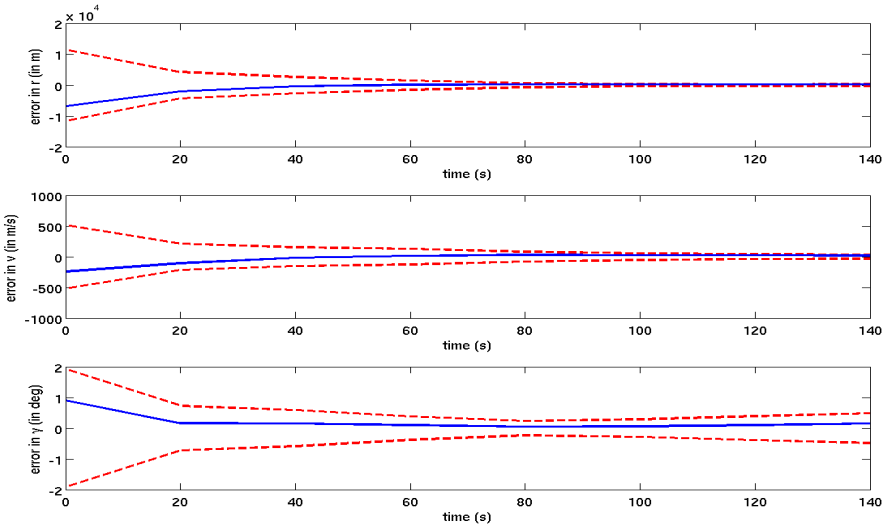


Fig. 28.: Generic particle filter with 7000 particles. True initial states are $[R_m +$ 61 km 2.64 km/s $-8.1^o]^T$ and update interval is 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

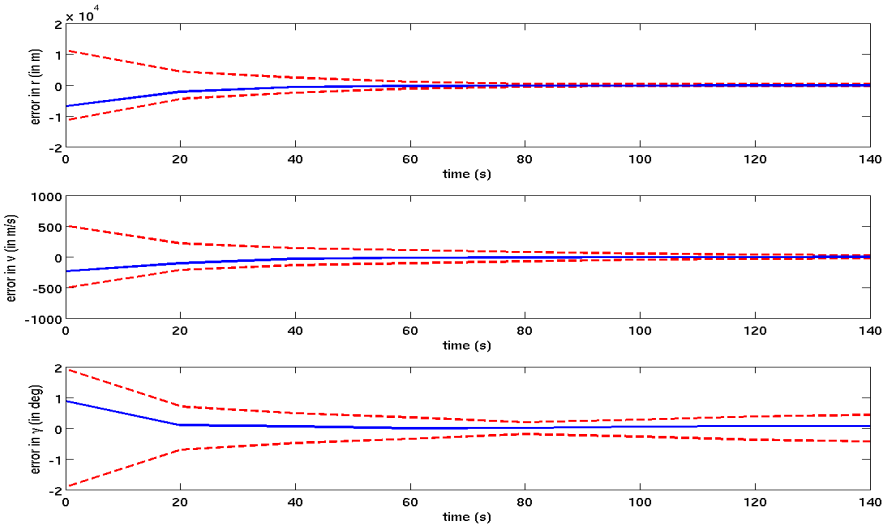


Fig. 29.: Bootstrap filter with 7000 particles. True initial states are $[R_m +$ 61 km 2.64 km/s $-8.1^o]^T$ and update interval is 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

The particle and the bootstrap filter did not perform as well as the Frobenius-Perron filter with lesser number of particles. Figure 28 and Fig. 29 show plots for error covariance and estimation errors for generic particle filter and bootstrap filter respectively with 7000 particle points. It can bee seen that they do not perform as well as FP operator based estimator (Fig. 26). In Fig. 30 and Fig. 31 the number of particles were increased to 10,000 for the particle filter based estimators. Even then, the performance of the particle filters do not match with FP operator based estimator's performance. To substantiate our claim, percentage error vs. time of the three estimators has been compared here. Figure 32 shows plots when the number of particles for each estimator is fixed at 7000. It can be seen that the percentage error in state estimate for FP operator based estimator is lower than the particle filter based estimators. The performance of the particle filters matches the FP operator only when the number of particles for generic particle filter is 25000 and for bootstrap filter is 20000, the percentage error plots for which is shown in Fig. 33.

Thus, it can be said that the Frobenius-Perron filter can achieve the same result as particle filters using significantly lower number of particles. These simulations were performed in a Linux machine with Intel® Pentium D processor. The superiority of the Frobenius-Perron filter is in the computational time. For each filtering algorithm, the computational time taken for one filtering step is presented in Table VI. It can be seen that due to fewer number of particles the Frobenius-Perron filter takes significantly less computational time.

The FP operator based estimator captures the non-Gaussian PDF evolution almost accurately. To show this, time evolution of 3^{rd} and 4^{th} moments of the posterior PDF for the FP operator based estimator with 7000 particles and a particle filter with large number of particles (100,000) has been plotted in Fig. 34. It can be seen that moments for FP operator matches exactly with the particle filter. If, the generic

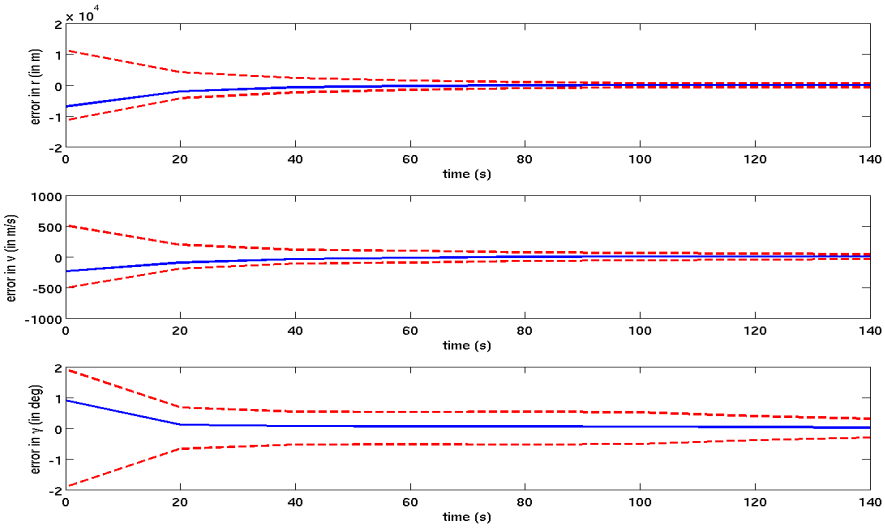


Fig. 30.: Generic particle filter with 10000 particles. True initial states are $[R_m +$ 61 km 2.64 km/s $-8.1^o]^T$ and update interval is 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

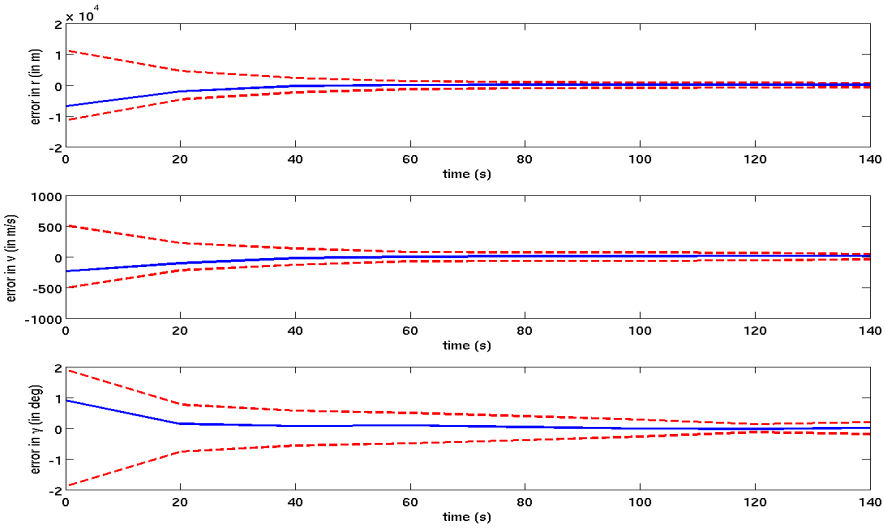


Fig. 31.: Bootstrap filter with 10000 particles. True initial states are $[R_m +$ 61 km 2.64 km/s $-8.1^o]^T$ and update interval is 20s. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

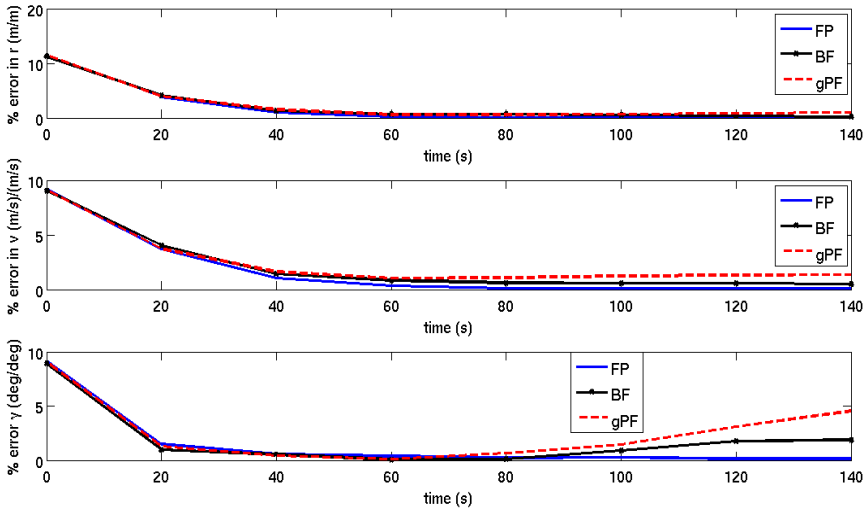


Fig. 32.: Percentage error in estimation with 7000 particles. True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$ and update interval is 20s. In the legend, 'BF', 'gPF' and 'FP', represent bootstrap filter, generic particle filter and Frobenius-Perron operator based estimator respectively.

Table VI.: Computational time taken per filtering step for each estimation algorithm. Listed times are in seconds

FP operator based filter	**Generic particle filter**	**Bootstrap filter**
57.42 s	207.96 s	168.06 s

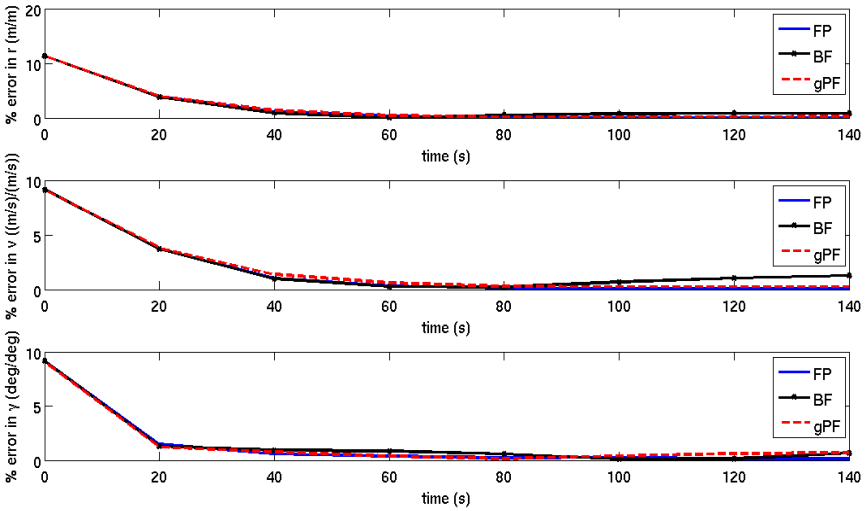
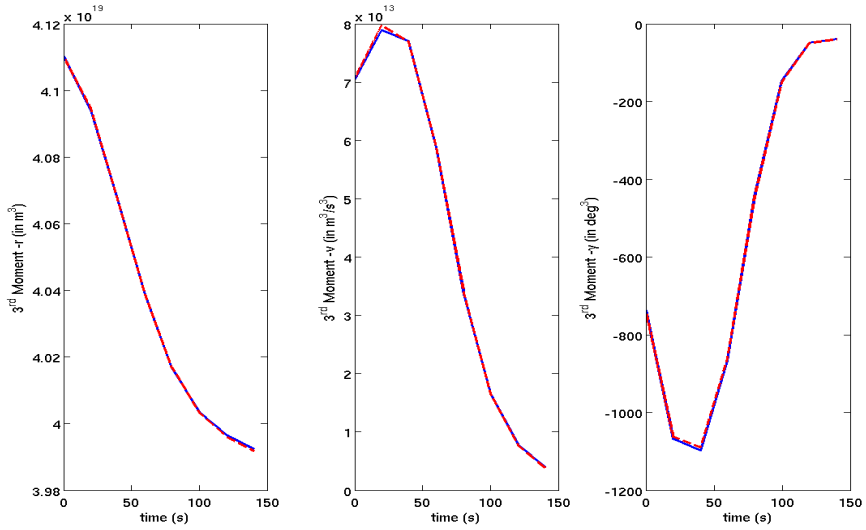
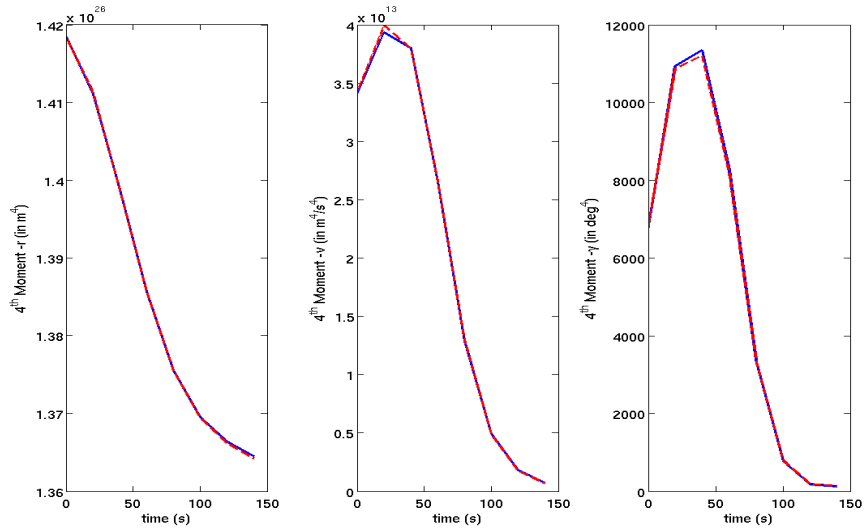


Fig. 33.: Percentage error in estimation with number of particles for FP based operator (FP), generic particle filter (gPF) and bootstrap filter (BF), 7000, 20000, and 25000 respectively. True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$ and update interval is 20s.

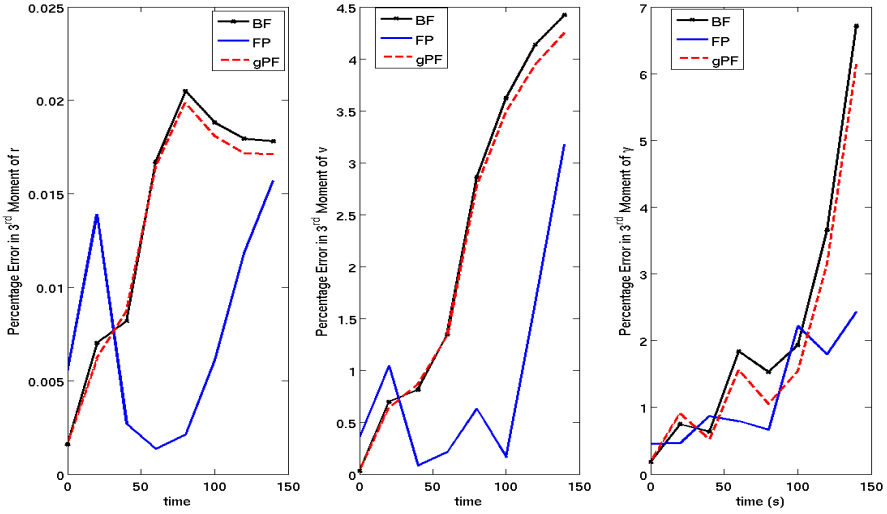


(a) 3rd order moments.

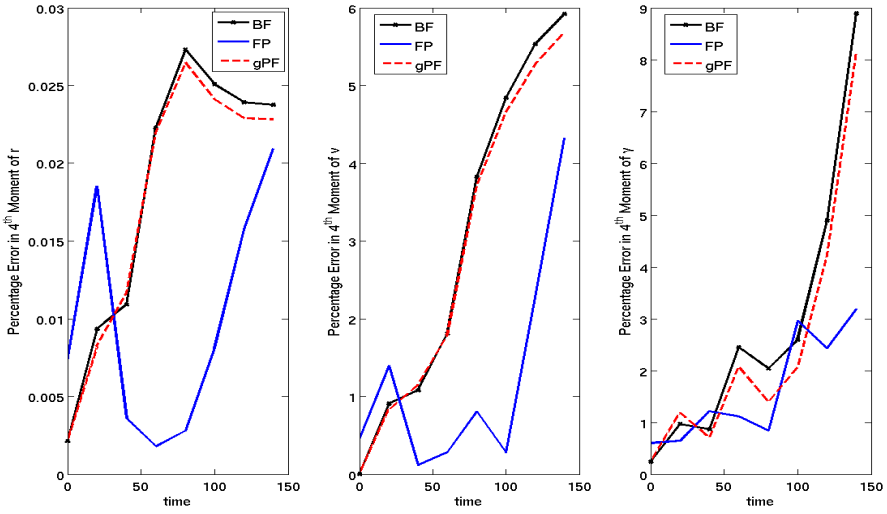


(b) 4th order moments.

Fig. 34.: Plots for 3rd and 4th order moments for particle filter with 100,000 particles- (dashed line) and FP operator based filter with 7000 samples- (solid line).



(a) Percentage error in 3rd order moments.



(b) Percentage error in 4th order moments.

Fig. 35.: Plots for percentage error in 3rd and 4th order moments for bootstrap filter (BF), generic particle filter (gPF) and FP operator based filter (FP), all with 7000 particles. Percentage deviation taken from particle filter with 100,000 particles.

particle filter and bootstrap filter with 7000 particles are compared, it is observed that they fail to predict non-Gaussian PDF evolution as accurately as the FP operator based estimator. To show this percentage deviation of the 3^{rd} and 4^{th} order moments of three estimators from the particle filter with 100,000 particles has been plotted in Fig. 35. It can be seen that percentage error for the FP operator based estimator lower than the particle filter based estimators.

a. Sensitivity Analysis

In this section, we perform sensitivity analysis of the FP operator based filter and the particle filters with respect to measurement noise and measurement update rate. Figure 36 and Fig. 37 show plots of error covariance and state estimate error for the FP operator based filter when the measurement noise is $6 \times 10^{-4}\mathcal{I}_3$ in scaled units, and when the measurement update rate is 40 seconds, respectively. The number of particles taken here is 7000. To achieve similar performance the generic particle filter and the bootstrap filter needed 20000 and 25000 particles respectively. The plots for particle filter are shown in Fig. 38 and Fig. 39 and for the bootstrap filter in Fig. 40 and Fig. 41.

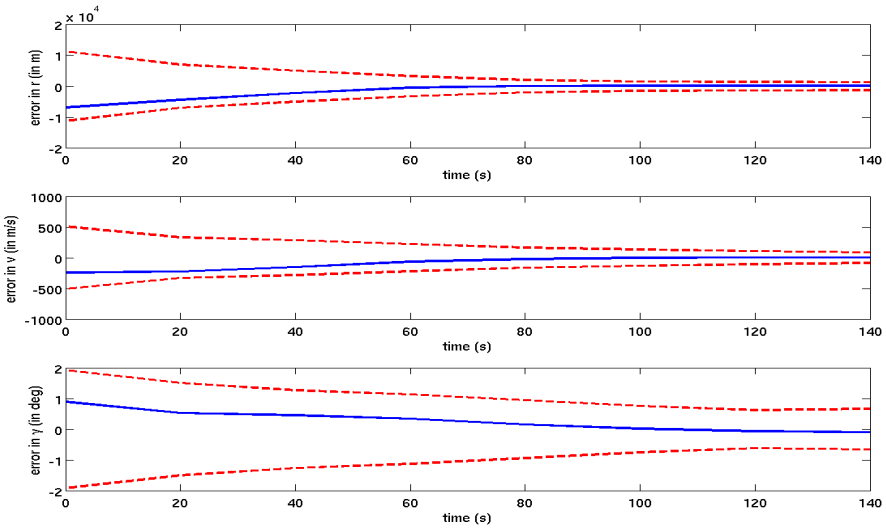


Fig. 36.: FP operator based filter with measurement noise $6 \times 10^{-4}\mathcal{I}_3$ in scaled units (number of samples = 7000). True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

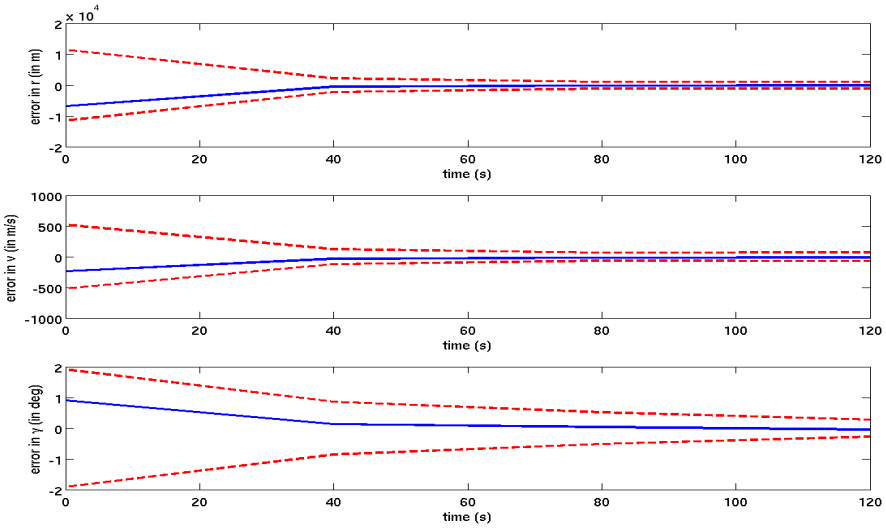


Fig. 37.: FP operator based filter measurement update interval is 40s (number of samples = 7000). True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

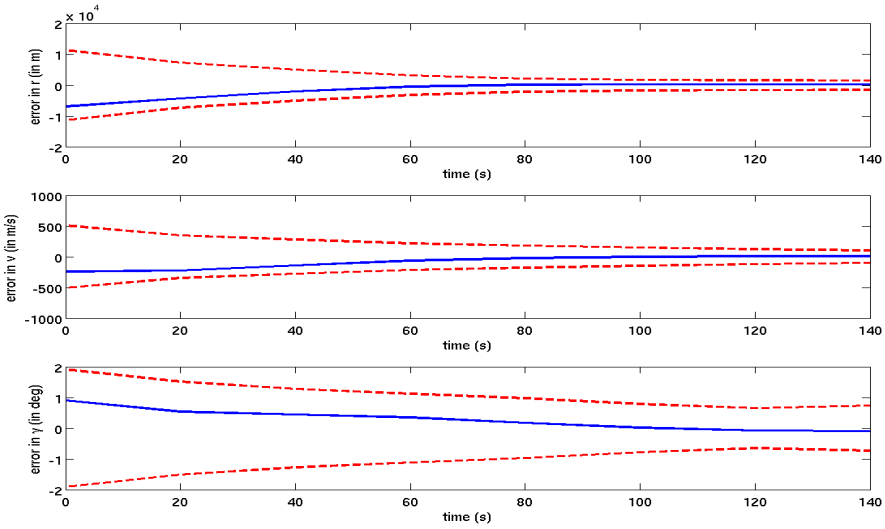


Fig. 38.: Generic particle filter with measurement noise $6 \times 10^{-4}\mathcal{I}_3$ in scaled units (number of samples = 25000) True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

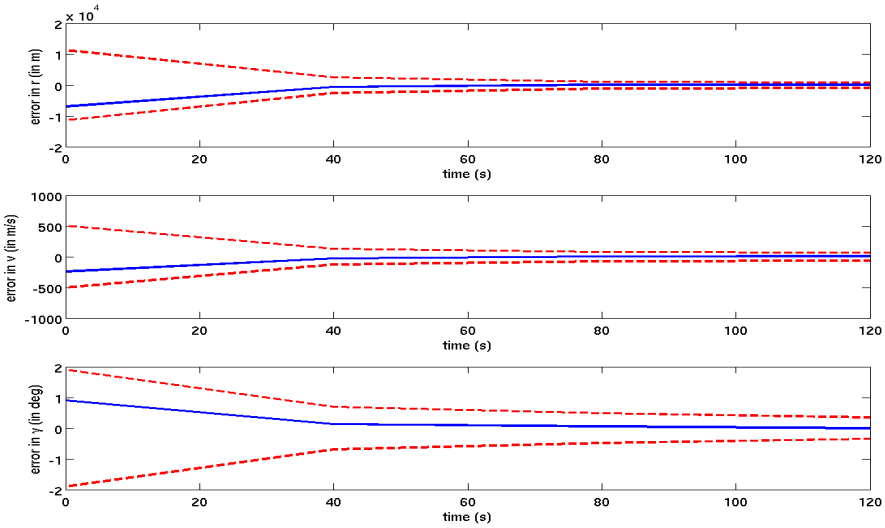


Fig. 39.: Generic particle filter measurement update interval is 40s (number of samples = 25000). True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

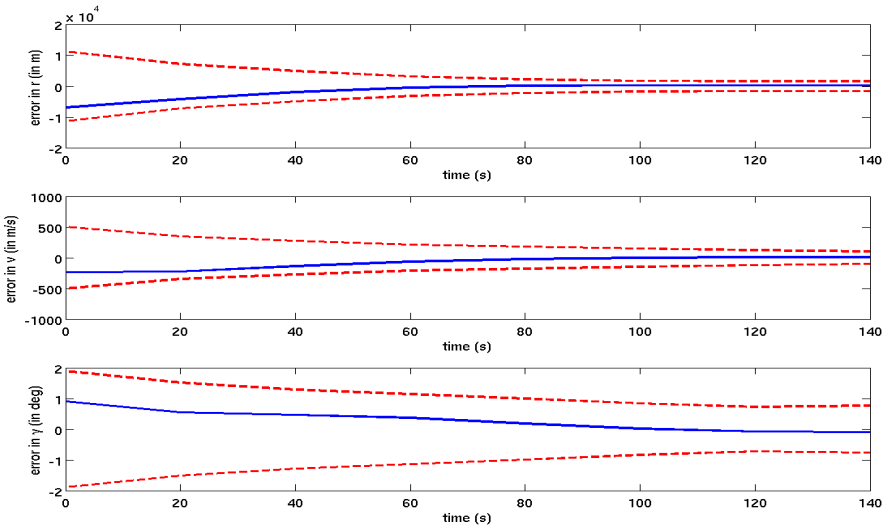


Fig. 40.: Bootstrap filter with measurement noise $6 \times 10^{-4}\mathcal{I}_3$ in scaled units (number of samples = 20000) True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

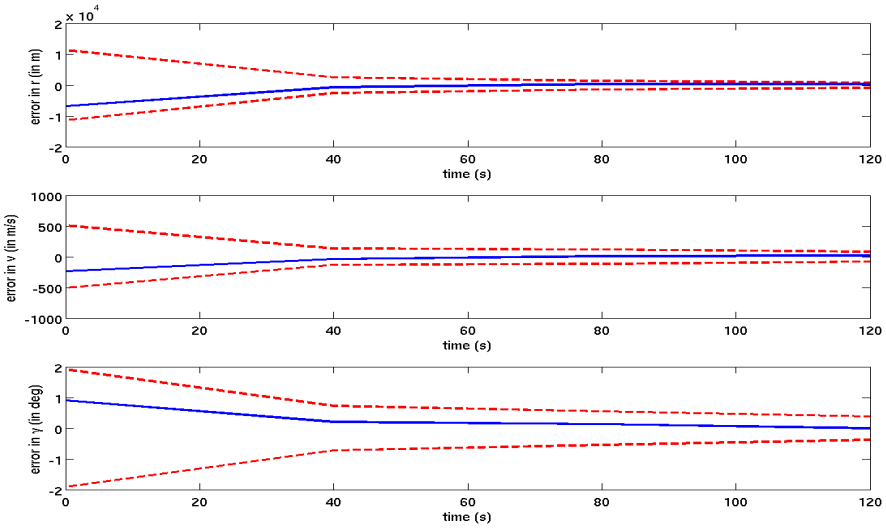


Fig. 41.: Bootstrap filter measurement update interval is 40s (number of samples = 20000). True initial states are $[R_m + 61 \text{ km } 2.64 \text{ km/s } -8.1^o]^T$. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

3. Application to Six-State Vinh's Equation

The estimation algorithm presented, is then applied to reentry of a hypersonic vehicle in Mars' atmosphere, with dynamics governed by six-state Vinh's equation [101]. The equation of motion are given by,

$$\dot{r} = v\sin\gamma, \tag{4.18a}$$

$$\dot{\theta} = \frac{v\cos\gamma\cos\xi}{r\cos\lambda}, \tag{4.18b}$$

$$\dot{\lambda} = \frac{v\cos\gamma\sin\xi}{r}, \tag{4.18c}$$

$$\dot{v} = -\frac{\rho v^2}{2B_c} - g\sin\gamma - \Omega^2 r\cos\lambda(\sin\gamma\cos\lambda - \cos\gamma\sin\lambda\sin\xi), \tag{4.18d}$$

$$\dot{\gamma} = \left(\frac{v}{r} - \frac{g}{v}\right)\cos(\gamma) + \frac{\rho}{2B_c}\left(\frac{L}{D}\right)v\cos\sigma + 2\Omega\cos\lambda\cos\xi + \frac{\Omega^2 r}{v}\cos\lambda(\cos\gamma\cos\lambda + \sin\gamma\sin\lambda\sin\xi), \tag{4.18e}$$

$$\dot{\xi} = \frac{\rho}{2B_c}\left(\frac{L}{D}\right)v\sin\sigma - \frac{v}{r}\cos\gamma\cos\xi\tan\lambda + 2\Omega(\tan\gamma\cos\lambda\sin\xi - \sin\lambda) - \frac{\Omega^2 r}{v\cos\gamma}\sin\lambda\cos\lambda\cos\xi, \tag{4.18f}$$

where θ is the geocentric longitude, λ is the geocentric latitude and ξ is the hypersonic vehicle's heading angle. Ω is the angular velocity of Mars given by 7.0882×10^{-5} rad/s, and σ is the bank angle which is taken as 0^o in this case.

The measurement model, $\tilde{y}$, consists measurements of geocentric latitude and longitude, θ and λ, along with dynamic pressure $\bar{q}$, heating rate H, and flight path angle, γ, whose expressions are given in Eqn. (3.35a) through Eqn. (3.35c). Hence, the measurement model is given by $\tilde{y} = [\bar{q}\ H\ G\ \theta\ \lambda]^T$.

The vehicle's heading angle ξ, is assumed to have no initial state uncertainty, and has initial value of 0.0573^o. Gaussian initial condition uncertainty, along $r, \theta, \lambda, v, \&\gamma$

is considered, with mean, μ_0 and standard deviation σ_0 given by

$$\mu_0 = [R_m + 54 \text{ Km}, -60^o, 30^o, 2.4 \text{ Km/s}, -9^o]^T \tag{4.19a}$$

$$\sigma_0 = \begin{pmatrix} 5.4 \text{ Km} & 0 & 0 & 0 & 0 \\ 0 & 3^o & 0 & 0 & 0 \\ 0 & 0 & 3^o & 0 & 0 \\ 0 & 0 & 0 & 240 \text{ m/s} & 0 \\ 0 & 0 & 0 & 0 & 0.9^o \end{pmatrix}. \tag{4.19b}$$

A nondimensionalized system is created, scaling each constants in Eqn. (4.18a) through Eqn. (4.18f), using scaling constants given in Table IV. The normalization factors used for measurements are given in Table VII.

Table VII.: Normalization factors used for measurements

Measurement	**Normalization Factors**
Dynamic Pressure	$1.97 \times 10^3 \frac{\text{N}}{\text{m}}$
Heating Rate	$0.0231 \frac{\text{J}}{\text{m-s}}$
Flight Path Angle	19.13^o
Geocentric Longitude	60.6^o
Geocentric Latitude	30.3^o

The performance of FP operator based filter is then compared with generic particle filter and bootstrap filter, when applied to the six-state Vinh's equation model. The initial states of the actual system are assumed to be,

$$[R_m + 61 \text{ km}, -60.6^o, 30.3^o, 2.42\text{Km/s}, -9.09^o, 0.0573^o]$$

in this case, with θ, λ, v, and γ having 1% error in initial state estimate. The

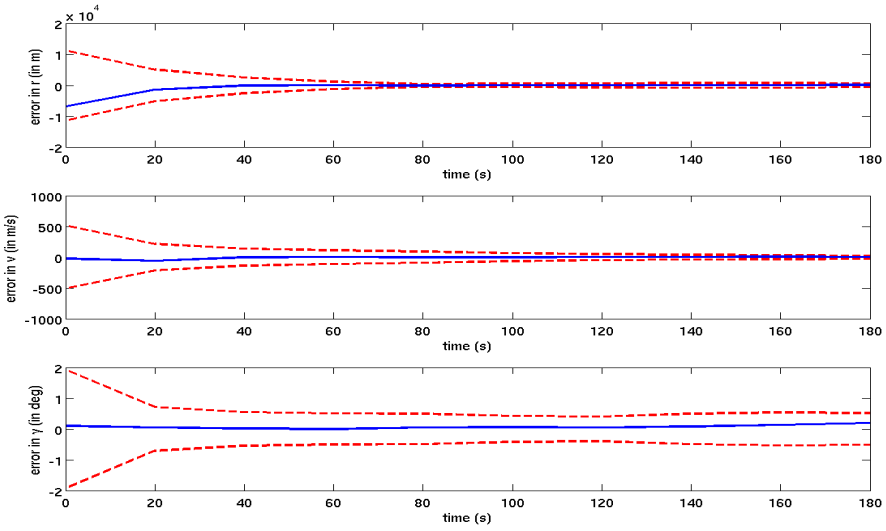
measurement update interval is kept fixed at 20 seconds. Figure 42 shows the plot for FP operator based filter with 9000 samples. Figure 43 and Fig. 44 shows plots for generic particle filter and bootstrap filter respectively, with same number of particles. It can be seen that FP operator based estimator performs better than particle filter based estimators. The errors for the generic particle filter and bootstrap filter diverge and are not inside the $\pm 3\sigma$ limits. Hence, these filters are inconsistent, when same number of particles as FP operator are taken.

The best performance that was achieved using particle filter and bootstrap filter is shown in Fig. 45 and Fig. 46 respectively. The number of particles needed for generic particle filter is 30000 and bootstrap filter is 25000. For particle filter, it is observed that although the $\pm 3\sigma$ bounds converge, the estimation errors escape outside these limits. For the bootstrap filter, the errors and the $\pm 3\sigma$ limits converge, but for γ the $\pm 3\sigma$ bounds diverge towards the end of estimation period.

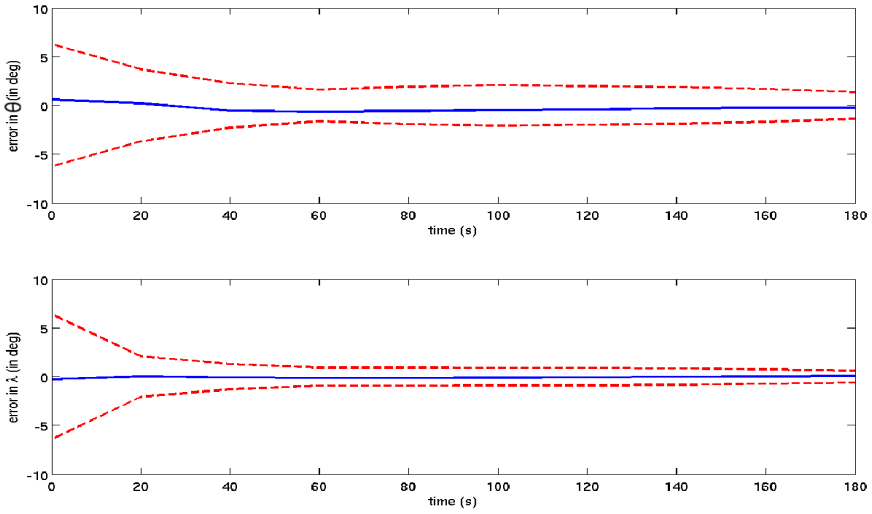
F. Limitations of Using Frobenius-Perron Operator

Although FP operator provides us with an exact methodology to propagate uncertainty in nonlinear dynamical systems, it has certain limitations. Throughout the analysis, we have dealt with systems that have only parametric or initial condition uncertainty. We have neglected the presence of process noise in a system.

For a system with process noise, evolution of densities are governed by the *Fokker Planck operator*. Thus Eqn. (4.7) consists of not only the first order, but second order derivatives too. The next chapter presents a complete treatise of such systems, and develops a methodology to propagate uncertainty in dynamical systems with process noise.

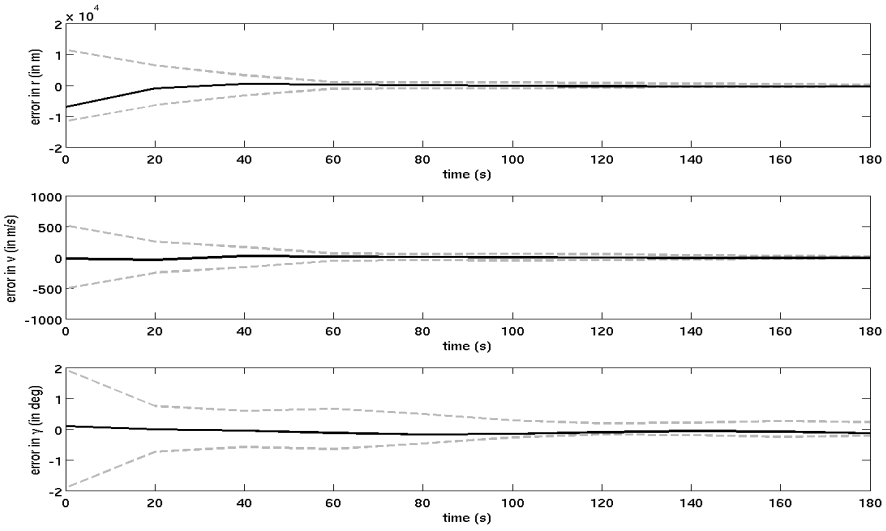


(a) Plots for r, v and γ.

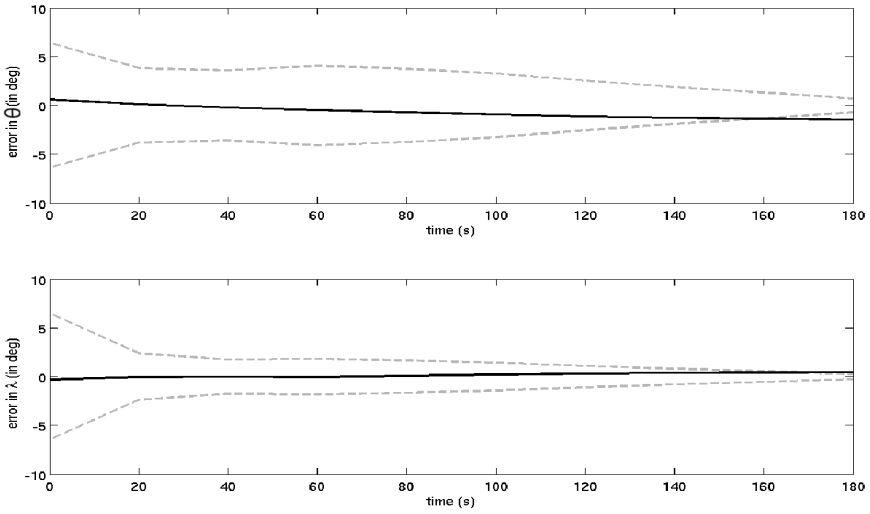


(b) Plots for θ and λ.

Fig. 42.: Performance of FP operator based filter with 9000 particles, when applied to six state Vinh's equation. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

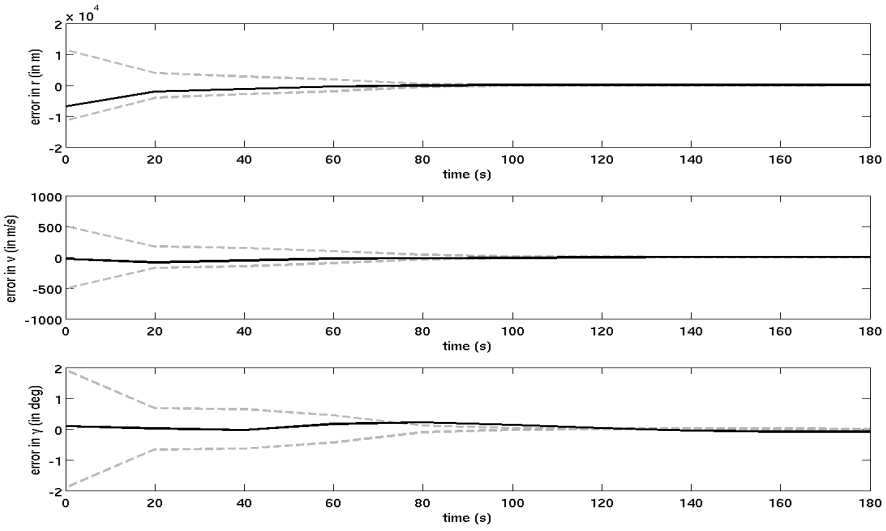


(a) Plots for r, v and γ.

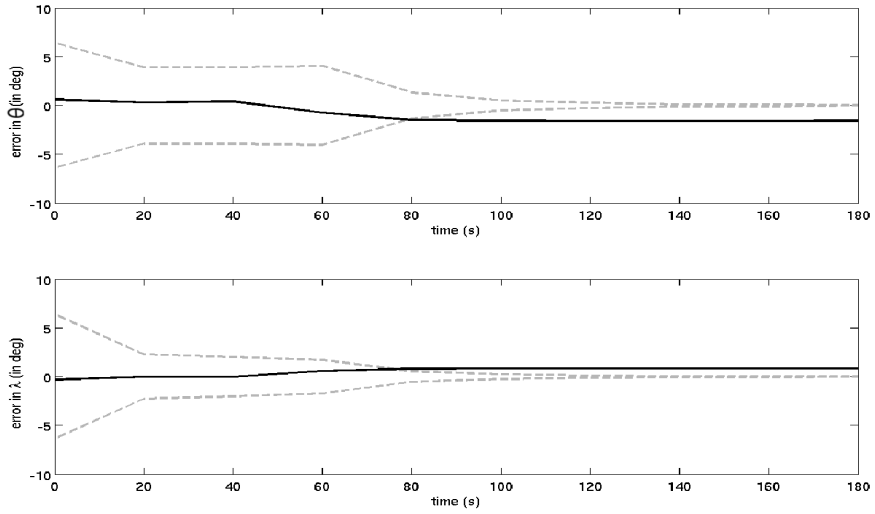


(b) Plots for θ and λ.

Fig. 43.: Performance of generic particle filter with 9000 particles, when applied to six state Vinh's equation. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

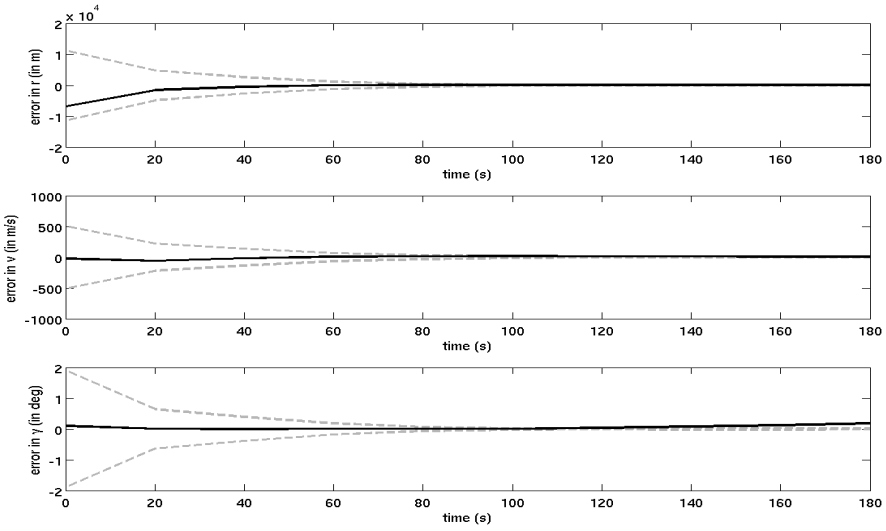


(a) Plots for r, v and γ.

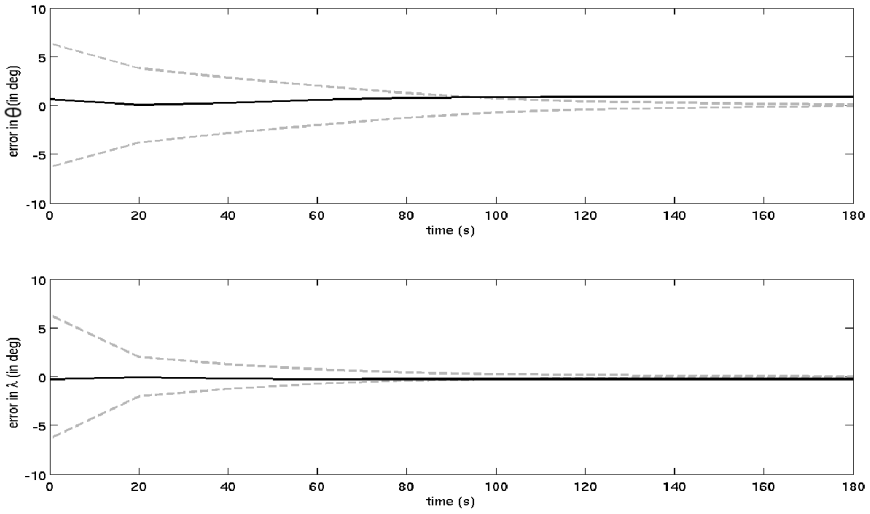


(b) Plots for θ and λ.

Fig. 44.: Performance of the bootstrap filter with 9000 particles, when applied to six state Vinh's equation. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

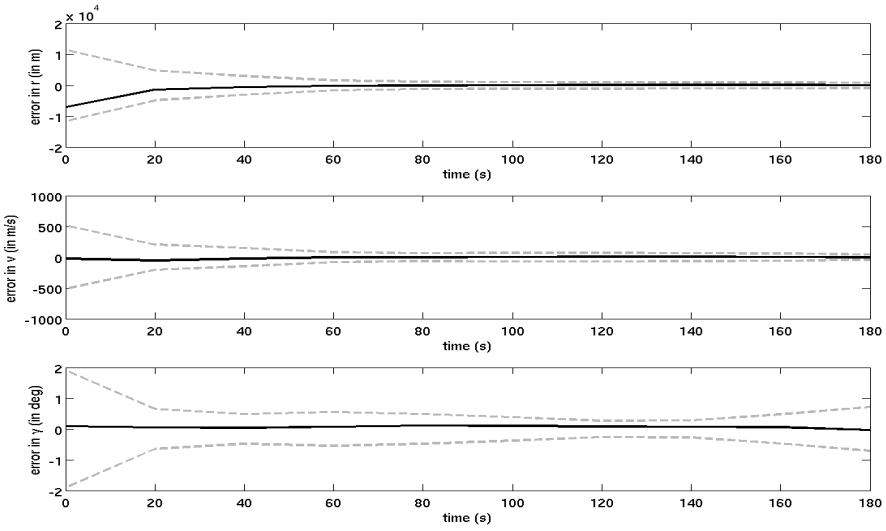


(a) Plots for r, v and γ.

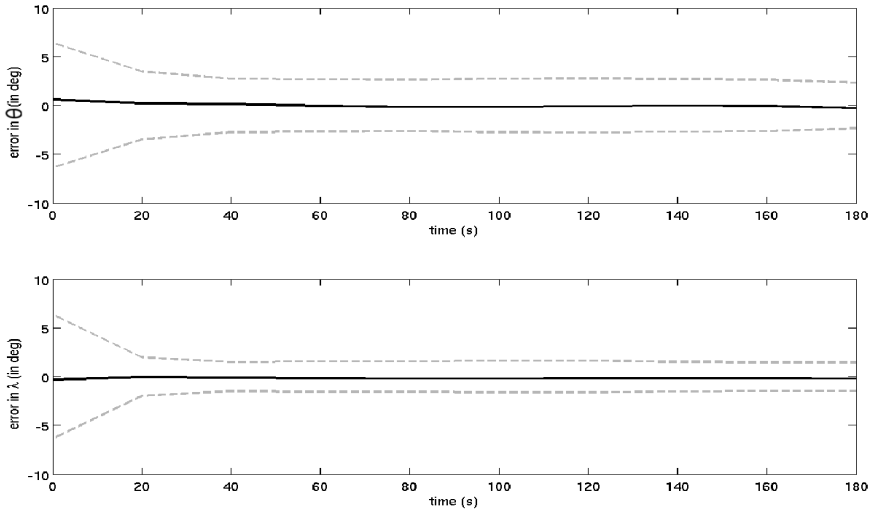


(b) Plots for θ and λ.

Fig. 45.: Performance of generic particle filter with 30000 particles, when applied to six state Vinh's equation. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.



(a) Plots for r, v and γ.



(b) Plots for θ and λ.

Fig. 46.: Performance of the bootstrap filter with 25000 particles, when applied to six state Vinh's equation. The dashed lines represent $\pm 3\sigma$ limits and the solid line represents error in estimation.

CHAPTER V

THE KARHUNEN LOÈVE EXPANSION

Up until this point, we have developed theories for uncertainty propagation in dynamical systems, which have only parametric or initial state uncertainty. However, the given model in a dynamical system may not match actual state evolution of the system. There may be uncertainties coming from unmodeled dynamics, modeling uncertainties and other phenomena which may be difficult to incorporate in the true system. Moreover, there may be some uncertainties in the actuation system, which was ignored during model development. Generally all the above mentioned uncertainties are coupled as one term, in the dynamical equation, which is called stochastic forcing term or the process noise term.

In this chapter, we will deal with the methodology applied to characterize the process noise term in an dynamical system. At first we will introduce such dynamical systems, methodology of solution, and how uncertainty propagates through such systems. Then, we will describe the *Karhunen Loève* (KL) expansion and how it can be used to represent the process noise term in a stochastic dynamical system, followed by some applications. Finally we will assess the convergence of the solution of the dynamical system, to the true solution, when the process noise term is expanded using a finite term KL expansion.

A. Stochastic Dynamical Systems

Consider the dynamical system given by,

$$dx\left(t\right) = f\left(x\left(t\right), \delta\right)\ dt + g\left(x\left(t\right), \delta\right)\ dW\left(t\right), \tag{5.1}$$

where $x(t) \in \mathbb{R}^n$ is the state vector at time t, $\delta \in \mathbb{R}^p$ is the vector of uncertain parameters, and $dW(t) \in \mathbb{R}^q$, is a random process denoting process noise. The function $f(.,.)$, $g(.,.)$ represent the dynamics while the function-valued matrix, $g(.,.)$ denotes noise coupling.

Equation (5.1) is referred to as stochastic dynamical system, or often stochastic differential equation. We will provide a brief overview of the solution methodologies of this equation, next.

1. Solution Methodology

Over the years, researchers have done a lot of work regarding solutions of stochastic differential equations [113, 114, 115, 116, 117]. A complete treatise of the solution methodology can be found in Øksendal [116]. Most commonly, $dW(t)$ is considered as a Wiener process [118] with $\mathbb{E}[W(t)] = 0$ and $\mathbb{E}[W(t)W(t')] = Q\delta(t - t')$ for all $t > t'$ with $t, t' \in \mathbb{R}^+$, where $Q \in \mathbb{R}^{q \times q}$ is the correlation matrix of the noise. Hence Eqn. (5.1) can be written as

$$\frac{dx(t)}{dt} = f(x(t), \delta) + g(x(t), \delta)\ \eta(t). \tag{5.2}$$

Here $\eta(t)$ is called the Gaussian white noise which is the derivative of the Wiener process. It has been seen [116], that there exists no process with continuous paths that satisfy $\eta(t)$. In other words the Gaussian white noise process is a nowhere differentiable process. Hence, solving Eqn. (5.2) is a hard problem. In the remainder of the section, we will try to avoid using Eqn. (5.2), but get some sense of the solution of the differential equation in Eqn. (5.1).

For general nonlinear systems, the state and the parametric uncertainty can be expressed as in an extended framework with states $z := [x, \delta]^T \in \mathbb{R}^{n+p}$. Without loss of generality, it is assumed that the dynamical system in Eqn. (5.1) has only state

uncertainty, hence $z = x$. The modified equation is given by

$$dx\left(t\right) = f\left(x\left(t\right), t\right)\ dt + g\left(x\left(t\right), t\right)\ dW\left(t\right). \tag{5.3}$$

Integrating Eqn. (5.3) till time T, we get

$$x(T) = x(0) + \int_0^T f\left(x\left(t\right), t\right)\ dt + \int_0^T g\left(x\left(t\right), t\right)\ dW\left(t\right). \tag{5.4}$$

Integral of the type $\int_0^T g\left(x\left(t\right), t\right)\ dW\left(t\right)$ is called the *Ito integral* [116]. It is difficult to solve this integral as the normal rules of Riemann-Stieltjes integration do not hold. Hence a distance preserving map is done to a space where it can be integrated.

Let $x(t) : (\Omega, \mathcal{F}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, given t, where $\mathcal{B}(\mathbb{R}^n)$ is the Borel σ-algebra on $\mathbb{R}^n$. Then the Ito integral can be written as $\int_0^T g\left(\omega, t\right)\ dW\left(t\right)$, where $\omega \in \Omega$.

Definition 3 *Let $V = V(S, T)$ be a class of function*

$$f(\omega, t) : \Omega \times [0, \infty) \to \mathbb{R},$$

such that

1. *$(\omega, t) \to f(\omega, t)$ is $\mathcal{F} \times \mathcal{B} -$ measurable, where $\mathcal{B}$ is the Borel σ-algebra on $[0, \infty)$.*
2. *$f(\omega, t)$ is $\mathcal{F}_t$-adapted.*
3. *$\mathbb{E}\left[\int_S^T f(\omega, t)^2 dt\right] < \infty$.*

Using *Ito Isometry* (Chapter 3 of [116]), it can be shown that

$$\mathbb{E}\left[\left(\int_S^T f\left(\omega, t\right)\ dW\left(t\right)\right)^2\right] = \mathbb{E}\left[\int_S^T f^2\left(\omega, t\right)\ dt\right] \forall f \in V(S, T). \tag{5.5}$$

It can also be shown under certain boundedness conditions, Eqn. (5.3) has unique solutions. Detailed discussion of this topic has been omitted here and can be found

in Øksendal [116].

2. Uncertainty Propagation in Stochastic Dynamical Systems

Given the stochastic dynamical system in Eqn. (5.3), let the initial state random variable be $x(0)$. Also let us assume that the states $x(t)$ admit a probability density function (PDF) $\rho(x,t)$, where $(x,t) \in \mathbb{R}^n \times [0,\infty]$, with the following properties.

1. *Positivity*: $\rho(x,t) > 0, \forall\ t\ \&\ x \in \mathbb{R}^n$.
2. *Normalization constraint*: $\int_{-\infty}^{\infty} \rho(x,t)dV = 1, \forall t \geq 0$ where $dV = dx_1 dx_2 \ldots dx_n$.

Further let the initial state PDF $\rho(x,0) = \rho_0$ and the PDF at $x = \infty$, $\rho(\infty,t) = 0$ be known. Then the evolution of uncertainty can be described by the following linear parabolic partial differential equation which is also called the *Fokker Planck Kolmogorov* (FPK) equation,

$$\frac{\partial \rho(x,t)}{\partial t} = -\sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(f_i \rho(x,t)\right) + \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j}\left(\left(gQg^T\right)_{ij} \rho(x,t)\right). \tag{5.6}$$

This equation is to be solved with the boundary conditions $\rho(0,x) = \rho_0$ and $\rho(t,\infty) = 0$. Solving the above partial differential equation is computationally hard [55, 38]. Researchers have tried to solve this problem either numerically [119, 120, 44, 43], or by using some basis function approximation [55, 45]. But all the solution methodologies suffer from curse of dimensionality [60].

But if there is no stochastic forcing term Eqn. (5.6) reduced to stochastic Liouville equation or Frobenius-Perron operator, whose solution methodology was presented in previous chapter. Here, we have used an approach of approximating the process noise term in Eqn. (5.3) using KL expansion. The advantage of such representation is that, the PDF evolution in the stochastic dynamical system can be found using the FP operator. In the following section, we describe the methodology using

KL expansion to represent noise in Eqn. (5.3) and solving for the PDF using FP operator.

B. The Karhunen Loève Expansion Applied to Stochastic Dynamical Systems

The Karhunen Loève (KL) expansion is a representation of a stochastic process as an infinite linear combination of orthogonal functions [61, 62]. The theory of KL expansion was developed from expanding any random process as a homogeneous product of functions of deterministic and the stochastic variable. This framework was first used in physics [63], to represent noise term in the *Langevin* equation. We will discuss about this specific application in subsequent sections. In this framework the class of such functions are chosen from the L_2 space, such that the they span the whole space and are linearly independent. Moreover, such a framework is advantageous as it helps us analyze the random process in an space where inner product is defined.

Given the random process $X(\omega, t)$ where $(\omega, t) \in \Omega \times [0, \infty)$, Ω being the sample space, let us define the class of functions $\phi_i(t) \in L_2$ satisfying the following properties.

1. $\langle \phi_i(t), \phi_j(t) \rangle = \delta_{ij}, \forall\ i, j = 1, 2, \ldots, \infty$, $\langle \cdot, \cdot \rangle$ represent inner product in L_2.
2. $\text{span}(\phi_1(t), \phi_2(t), ...) = \text{span}(L_2)$.

Due to the fact that the $\phi_i(t)$ form the basis in L_2 space, there are certain restrictions on the random processes that can be represented in such a framework. Here, we require that the random process is of second order, or $\mathbb{E}\left[X(\omega, t)^2\right] < \infty, \forall\ t \in [0, \infty)$, and that it is centered, i.e., $\mathbb{E}\left[X(\omega, t)\right] = 0, \forall\ t$.

Now, as we have specified the class of approximation functions to be used, we will define the homogeneous representation of any random process. It is given by,

$$X(\omega, t) = \sum_{i=1}^{\infty} \xi_i(\omega)\phi_i(t), \tag{5.7}$$

where $\xi(\omega)$ are random variables which are pairwise uncorrelated i.e. $\mathbb{E}\left[\xi_i(\omega)\xi_j(\omega)\right] = \mathbb{E}\left[\xi_i(\omega)\right]\mathbb{E}\left[\xi_j(\omega)\right], \forall\ i, j = 1, 2, \ldots, \infty$. Our aim here is to use a finite term approximation of Eqn. (5.7), and to assess convergence in some sense.

Given the covariance function of the random process $X(\omega, t)$ as $C_{XX}(t_1, t_2)$, with eigenvalues and eigenvectors as $\{\lambda_i\}_{i=1}^{\infty}$ and $\{\psi_i(t)\}_{i=1}^{\infty}$, the random process $X(\omega, t)$, can be expanded in a similar manner as Eqn. (5.7), in terms of the eigenvalues and eigenfunctions of its covariance function. It is given by,

$$X(\omega, t) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\xi_i(\omega)\psi_i(t) \tag{5.8}$$

where $\xi_i(\omega)$ are pairwise uncorrelated random variables. The expansion in Eqn. (5.8) is the representation of the Karhunen Loève expansion. It can be proved that a finite term KL expansion converges to the true random process uniformly in t and in mean square sense in ω [121]. In other words, if number of terms in the KL expansion is N, then

$$\mathbb{E}\left[X(\omega, t) - \sum_{i=1}^{N} \sqrt{\lambda_i}\xi_i(\omega)\psi_i(t)\right] \to 0, \text{ uniformly in } t \text{ as } N \to \infty.. \tag{5.9}$$

1. KL Expansion of Wiener Process

Let $W(\omega, t)$ be a Wiener process, then the covariance function is given by, $C_{WW}(t_1, t_2) = Q \times \min(t_1, t_2)$, where $Q \in \mathbb{R}^{q \times q}$, is the real valued matrix representing the autocorrelation of the process. Here, we have $t \in [0, T]$ where $T < \infty$, hence the eigenvalues and eigenvectors are given by,

$$\lambda_i = \frac{4T^2}{\pi^2\left(2i-1\right)^2}, \tag{5.10a}$$

$$\psi_i\left(t\right) = \sqrt{2}\ \sin\left(\left(i - \frac{1}{2}\right)\frac{\pi t}{T}\right). \tag{5.10b}$$

Hence,

$$W(\omega,t)=\sqrt{2}\sum_{i=1}^{\infty}\xi_i(\omega)\frac{\sin\left(i-\frac{1}{2}\right)\frac{t}{T}}{\left(i-\frac{1}{2}\right)\frac{t}{T}}, \tag{5.11}$$

where, $\xi_i(\omega)\sim\mathcal{N}(0,Q)$. The KL expansion of a Gaussian white noise can be found by differentiating Eqn. (5.11) with respect to t (page 548 of [122]). If $\eta(\omega,t)$ is a Gaussian white noise, then

$$\eta(\omega,t)=\sqrt{2}\sum_{i=1}^{\infty}\xi_i(\omega)\cos\left(i-\frac{1}{2}\right)\frac{t}{T} \tag{5.12}$$

The above expansion will be used next to represent noise in Eqn. (5.2).

2. KL Expansion Applied to Langevin Equation

Consider the stochastic dynamical system in Eqn. (5.2), with $g(x(t),\delta)=1$. We will incorporate the sample space of the Gaussian white noise $\omega\in\Omega$ in its argument. Important point to note is, Ω is independent of the sample space of $X(t)$. Without loss of generality, we consider only state uncertainty in the system. Therefore, Eqn. (5.2) can be written as,

$$\dot{x}(t)=f(x,t)+\eta(\omega,t) \tag{5.13}$$

Here we will consider integrating Eqn. (5.13) up to a finite time T, hence, $\eta(\omega,t):\Omega\times[0,T]\rightarrow\mathbb{R}^n$ is a Gaussian white noise having autocorrelation Q. The above equation is often referred to as Langevin equation [123]. Utilizing the framework described in the previous section, we use KL expansion of white noise given by Eqn. (5.12). Hence Eqn. (5.13) can be written as,

$$\dot{x}(t)=f(x,t)+\sqrt{2}\sum_{i=1}^{\infty}\xi_i(\omega)\cos\left(i-\frac{1}{2}\right)\frac{t}{T}. \tag{5.14}$$

Here we will use a finite term approximation of the Eqn. (5.14) and then assess the convergence of solutions of approximated system to the actual solution. Let the number of terms in the KL expansion in the RHS of Eqn. (5.14) be N, and the approximate solution of Eqn. (5.14) be given by $x_N(t)$. The approximated system is then given by,

$$\dot{x}_N(t) = f(x_N, t) + \sqrt{2}\sum_{i=1}^{N} \xi_i(\omega) \cos\left(i - \frac{1}{2}\right)\frac{t}{T}. \tag{5.15}$$

C. Convergence of Solutions

Due to finite term approximation of the process noise in Eqn. (5.13), the solution obtained after propagation is not error-free. It is well known that, KL expansion converges in mean square sense, to the underlying stochastic process as $N \to \infty$. But, same argument cannot be extended regarding the convergence of states $x_N(t)$ of the dynamical system in Eqn. (5.15). To obtain a meaningful solution, there should be some notion of convergence of the states in Eqn. (5.15) to states in Eqn. (5.13). Hence, verification of the obtained solution is important in this scenario. Here, first we will try to verify a weaker notion of convergence, before proving to a stronger condition on the solutions. We first verify that the solution of Eqn. (5.15) and Eqn. (5.13) converge in distribution. Then we have proved that under certain conditions, the solutions converge in mean square sense.

1. Verification of Convergence in Distribution

In this section we have verified that the solution of the approximated dynamical system converge to the true solution in distribution. Our approach here is mainly application specific. Therefore, we have verified that given a dynamical system whether the solution of the approximated system and the actual system converge in distribu-

tion. The application focused on here is a Vanderpol's oscillator. The methodology for verification used is statistical hypothesis testing [124].

We have used the Kolmogorov-Smirnov (KS) test to verify the solution of approximated dynamics. The Kolmogorov-Smirnov test [125] is a statistical test used to compare a sample with a reference probability distribution. It quantifies a distance D_M, between reference cumulative distribution function (CDF), $\pi(x)$, and empirical CDF of the sample being tested, $\pi_M(x)$, which is given by

$$D_M = \sup_x |\pi_M(x) - \pi(x)|. \tag{5.16}$$

Here, M refers to sample size of the given sample. The null hypothesis is that the sample comes from the reference distribution. Given a significance level α, the null hypothesis is accepted if,

$$\sqrt{M} D_M \leq K_\alpha, \tag{5.17}$$

where K_α can be found from ,

$$\Pr\left(K \leq K_\alpha\right) = 1 - \alpha. \tag{5.18}$$

Here, K is a random variable which follows Kolmogorov distribution, with CDF,

$$\Pr\left(K \leq x\right) = \frac{\sqrt{2\pi}}{x} \sum_{i=1}^{\infty} e^{-\frac{(2i-1)^2 \pi^2}{8x^2}}. \tag{5.19}$$

If Eqn. (5.17) is satisfied, then the KS test is passed and it can be concluded that the given sample comes from the reference CDF.

As the sample size M is increased the empirical measure π_M approaches the true measure π_0. Hence, if the sample is from π_0, D_M should go to zero. This can be

shown using the Glivenko-Cantelli lemma which states that,

$$\sup_{x\in\mathbb{R}} |\pi_M - \pi_0| \to 0 \qquad \text{almost surely.} \tag{5.20}$$

Due to the fact that $M < \infty$, D_M is a random variable, and each sample will give us different D_M values. The empirical distribution of $\sqrt{M}D_M$ is given by,

$$\pi_{M_s}(x) = \frac{1}{M_s}\sum_{i=1}^{M_s} \mathcal{I}_{x<\sqrt{M}D_M^i}, \tag{5.21}$$

where M_s is the number of samples and D_M^i is the D_M value for i^{th} sample. Dvoretzky Kiefer Wolfowitz inequality [126] states that,

$$\Pr\left(\sup_{x\in\mathbb{R}} |\pi_{M_s}(x) - \pi_K(x)| > \varepsilon\right) \leq 2e^{-2M_s\varepsilon^2} \ \forall \varepsilon > 0, \tag{5.22}$$

where $\pi_K(x)$ is the Kolmogorov CDF in Eqn. (5.19). Hence as $M_s \to \infty$ the test statistic $\sqrt{M}D_M$ exponentially converges in distribution to a Kolmogorov random variable with rate $2e^{-2M_s\varepsilon^2}$. Detailed discussion of KS test can be found in ref. [125, 127].

In the present case, the analytical representation of CDF at time $t = T$ is not known. Hence to verify our solution, we back-propagate the sample obtained after propagation of Eqn. (5.15) at final time T to time $t = 0$ using the original dynamics in Eqn. (5.13); and check if the back-propagated sample belongs to the initial CDF. For simplicity of comparing CDFs we have used inverse transform sampling [128] to convert the initial samples and samples obtained after back-propagation to uniformly distributed samples in $[0, 1]$.

Let the initial state, $x(0) = x_0 \in \mathbf{D}_0$, have a CDF π_0. Let the elements of the sample; sampled from the initial CDF be $x_{0,i}, i = 1, \ldots, M$, and elements after propagation of Eqn. (5.15) be $x_{T,i}$. Let the back-propagated sample obtained by

propagating Eqn. (5.13) from $[T, 0]$ be $\hat{x}_{0,i}$. Here, we claim that, $\hat{x}_{0,i}$ is a sample from the initial CDF. Hence, using the KS test, the sample $\hat{x}_{0,i}$, is compared with the initial CDF, for statistical significance. Here, we have used a multivariable KS test, which is similar to the one presented in [125]. The verification methodology employed, is described in algorithm 1.

We apply the proposed verification methodology to verify states of the Vanderpol's oscillator given by,

$$\dot{x}_1(t) = x_2(t) \tag{5.23a}$$

$$\dot{x}_2(t) = \left(1 - x_1^2(t)\right) x_2(t) - x_1(t) + \eta(\omega, t). \tag{5.23b}$$

We assume, the initial conditions are normally distributed, with CDF given by, $\pi_0 \sim \mathcal{N}([0, 0], \text{diag}[1, 1])$, and the process noise has the autocorrelation $Q = 2\pi$. We obtain $M_s = 100$ samples, each of sample size $M = 500$ from π_0. The number of terms in the KL expansion is fixed to $N = 21$. At first, we pick a sample from the 100 available samples and perform KS test on it. Figure 47 shows plot for the location of elements initially (x_0) and after back-propagation $(\hat{x}_0)$, for the particular sample. It is observed that the back-propagated sample, is clustered around the origin and sparseness increases as we move away. This is in agreement with the physical intuition of random samples drawn from π_0 which is a standard normal distribution.

Using the methodology in [125] we obtain uniformly distributed sample $\hat{y}_0 = F(\hat{x}_0)$ using inverse transform sampling theory [128]. Figure 48 shows the plot of empirical CDF of $\hat{y}_0^j$ and the uniform CDF for the selected sample. It can be seen that, visually the CDFs are close to each other. Hence, the value of D_M for the particular sample is expected to be less.

Algorithm 1 Verification of solutions of approximated dynamical systems using KL expansion

Require: Domain $\mathbf{D}_0$, CDF π_0, of initial parameters, significance level α, sample size M, number of samples M_s, number of terms in KL expansion N.

1: Calculate K_α using Eqn. (5.17).

2: **for** $j = 1$ to M_s **do**. ▷ Sample number counter.

3: Draw M elements $x^j_{0,i}$ at random from $\mathbf{D}_0$. ▷ Use Markov Chain Monte Carlo.

4: Create null hypothesis $H_0 : \pi = \pi_0$ for the current sample

5: **for** $i = 1$ to M **do**. ▷ Sample element counter.

6: Propagate $x^j_{0,i}$, from $t = [0, T]$, using Eqn. (5.15) to get $x^j_{T,i}$. ▷ Use of KL approximated dynamics with number of terms N.

7: Backward propagate $x^j_{T,i}$, from $t = [T, 0]$ using Eqn. (5.13) to get $\hat{x}^j_{0,i}$. ▷ Use of original stochastic dynamics.

8: Get $\hat{y}^j_{0,i} \leftarrow \pi_0(\hat{x}^j_{0,i})$ ▷ From inverse transform sampling theory, $\hat{y}^j_0$ is uniformly distributed in $[0, 1]$.

9: **end for**

10: Calculate empirical CDF of the current sample $\hat{y}^j_0$, $G^j_M(y) \leftarrow \frac{1}{M} \sum_{i=1}^{M} \mathcal{I}_{y \le \hat{y}^j_{0,i}}$.

11: Calculate $D^j_M \leftarrow \sup_{y \in [0,1]} |G^j_M(y) - G(y)|$. ▷ $G(y)$ is the uniform CDF in $[0, 1]$.

12: **if** $D^j_M \le \frac{K_\alpha}{\sqrt{M}}$ **then**.

13: Accept H_0 for the j^{th} sample.

14: **else**

15: Reject H_0 for the j^{th} sample.

16: **end if**

17: **end for** ▷ Repeat the same test for the next sample

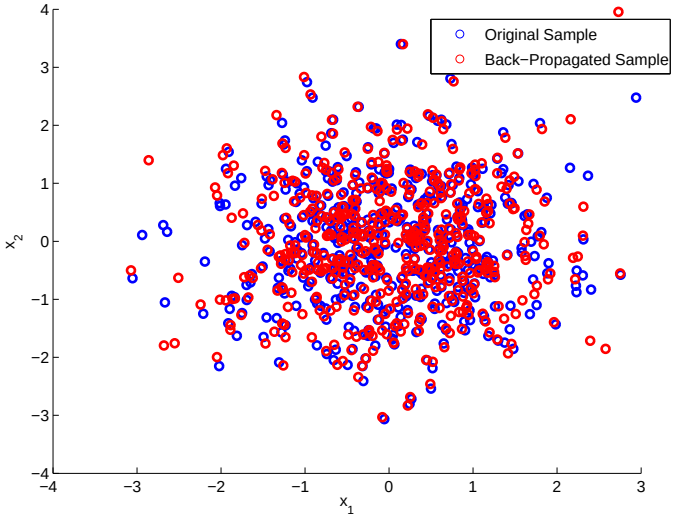


Fig. 47.: Scatter plot of initial and the back-propagated elements for a given sample. Blue circles are the original sample and the red ones are the back-propagated sample.

For the KS test, the value of α is fixed to 0.05. Hence, using Eqn. (5.18) and Eqn. (5.19), we get $K_\alpha/\sqrt{M} = 0.0607$. For the given sample, the value of D_M was 0.0455. Hence, the null hypothesis, that the sample is drawn from a standard normal distribution, is accepted with a significance level $\alpha = 0.05$. Figure 49 shows plots for D_M values for all the 100 samples drawn. The samples above red line fail the test if $\alpha = 0.05$ and the samples above the black line fail, if $\alpha = 0.01$. It is observed that 23 samples fail the test when $\alpha = 0.05$ and 5 samples fail when $\alpha = 0.01$. It can be seen that, as the significance level α, is decreased, the chances that the null hypothesis is accepted increases.

To get an idea of rate of convergence in solution with number of terms in KL expansion, N, we plot the D_M values by increasing N. Figure 50 shows plot for D_M values for 76 and 91 terms in KL expansion. It was found that with an $\alpha = 0.05$,

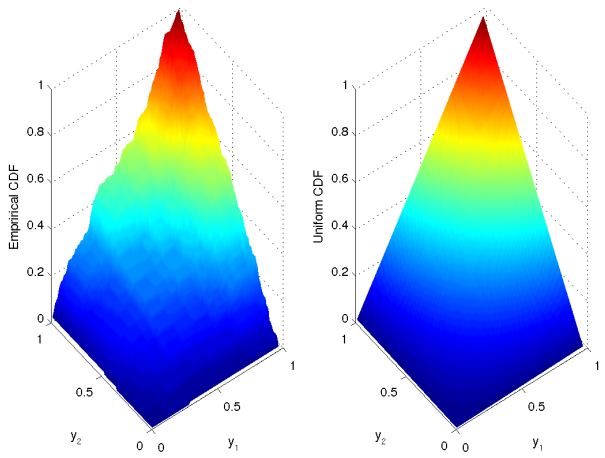


Fig. 48.: Plot of empirical CDF of $\hat{y}_0^j$ and uniform CDF for a given sample.

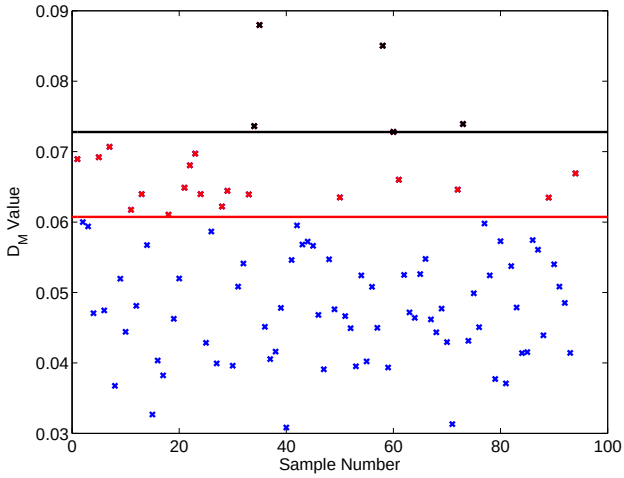
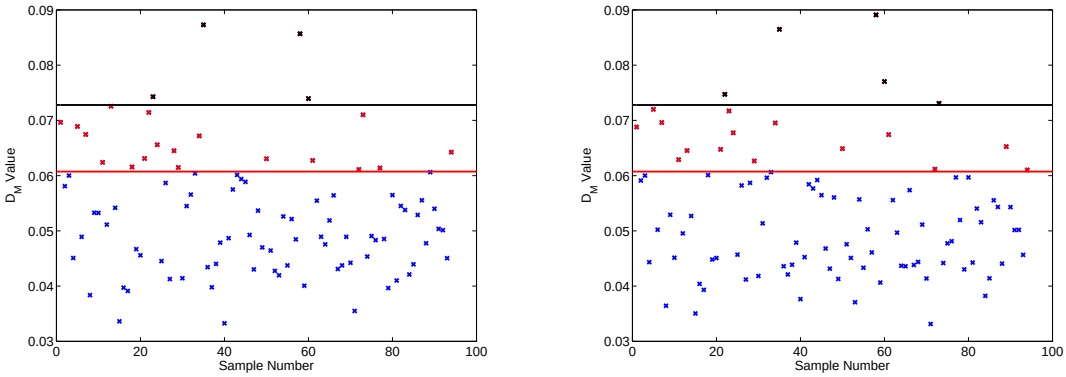


Fig. 49.: Plot D_M value for all the 100 samples. The samples above the red and black lines fail when $\alpha = 0.05$ and 0.01 respectively.

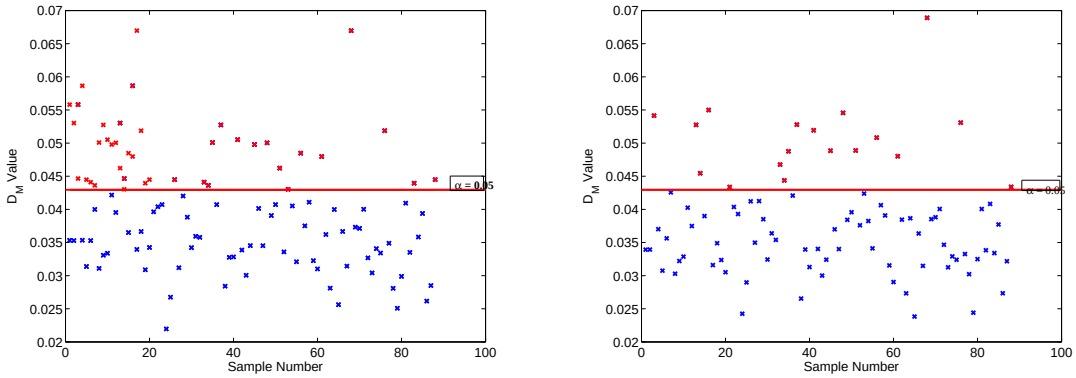


(a) $N = 76$. (b) $N = 91$.

Fig. 50.: Plot D_M value for all the 100 samples for a. $N = 76$ and b. $N = 91$. The samples above the red and black lines fail when $\alpha = 0.05$ and 0.01 respectively.

22 samples fail the test when $N = 76$ and 20 samples fail with $N = 91$. With an $\alpha = 0.01$, number of samples failing the test for $N = 76, 91$ are 4 and 5 respectively. Next, we vary M and plot the D_M values for variation in M. Figure 51 shows plots for $M = 1000$ with $N = 21$ and 76. It is observed that number of samples failing the KS test with $\alpha = 0.05$ is 20 for $N = 21$ and 18 for $N = 76$. Hence we can conclude that, there is some sense of convergence in solution, though for the particular application, the convergence is very slow.

Now we will verify, the consistency of the test statistic D_M, i.e. if it follows the prescribed distribution and assess its convergence. As mentioned before in Eqn. (5.22), the value of D_M obtained after KS test is a random variable, whose CDF converges to Kolmogorov distribution exponentially. Figure 52 shows plots for the empirical CDF of D_M given by Eqn. (5.21) and the Kolmogorov CDF in $\mathbb{R}[0, 1]$, for a fixed $N = 21$. We vary the the number of samples, M_s and observe the convergence



(a) $N = 21$. (b) $N = 76$.

Fig. 51.: Plot D_M value for all the 100 samples for a. $N = 21$ and b. $N = 76$, with increased $M = 1000$. The samples above the red line fails when $\alpha = 0.05$.

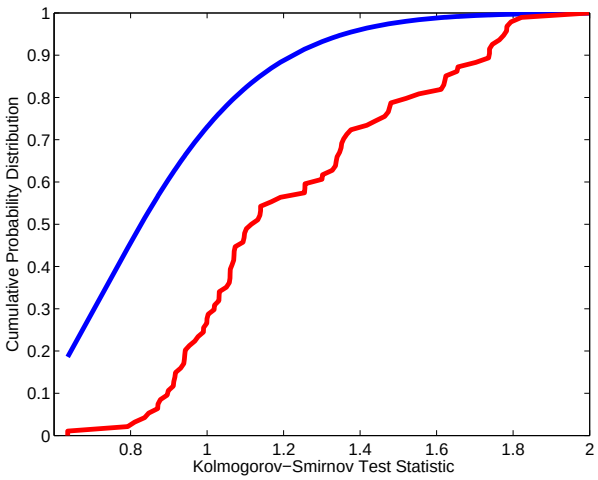


Fig. 52.: Plot of empirical CDF of $\sqrt{M}D_M$ (red) and Kolmogorov CDF (blue).

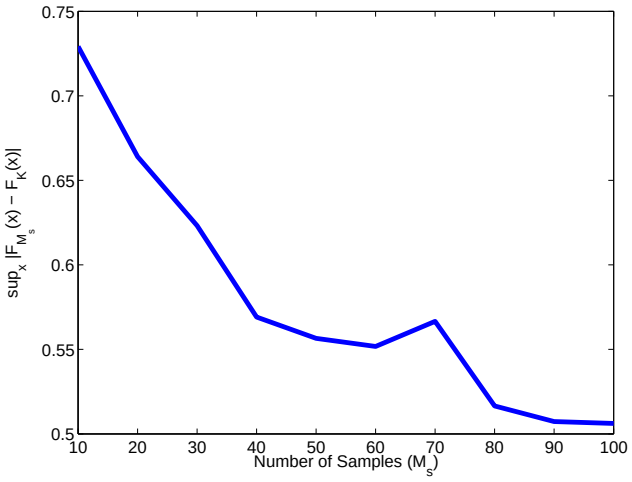


Fig. 53.: Plot of $\sup_{x\in\mathbb{R}} |F_{M_s}(x) - F_K(x)|$ vs. M_s .

in $\sup_{x\in\mathbb{R}} |F_{M_s}(x) - F_K(x)|$. The variation is plotted in Fig. 53. It can be seen that $\sup_{x\in\mathbb{R}} |F_{M_s}(x) - F_K(x)|$ decreases exponentially as we increase M_s, which is in compliance with the theory. Due to the fact that the convergence is in probability but only one realization of $\sup_{x\in\mathbb{R}} |F_{M_s}(x) - F_K(x)|$ has been plotted here, we do not observe a monotonic behavior in convergence. But, there are very few outliers, as the probability in the left hand side of Eqn. (5.22) can be decreased arbitrarily by varying ϵ.

2. Convergence in Mean Square Sense

In the last section, we have shown that there is some sense of convergence between solutions of the approximated system in Eqn. (5.15) and the actual system in Eqn. (5.13). We will here prove a stronger notion of convergence that the solutions actually converge in mean square sense. Before, showing that we state the following theorem,

Theorem 1 *Let $x(\omega, t)$ be the solution of the nonlinear stochastic differential equa-*

tion (SDE),

$$dx(t) = f(x,t)dt + dW(\omega,t) \tag{5.24a}$$

$$\Rightarrow \frac{dx\,(t)}{dt} = f\,(x,t) + \eta(\omega,t), \tag{5.24b}$$

where $f : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ *satisfies the following conditions*

1. *Non-explosion condition: there exists* $D \geq 0$, $|f\,(x,t)| < D(1+|x|)$ *where* $x \in \mathbb{R}^n$.
2. *Lipschitz condition: there exists* $C \geq 0$, *and* $x, y \in \mathbb{R}^n$ $|f\,(x,t) - f\,(y,t)| < C\,|x-y|$.

Let $x_N\,(t)$ *be solution of the differential equation, formed by* $N-$*term approximation of* $\eta(\omega,t)$ *in Eqn. (5.24b), using orthonormal basis, which is given by*

$$\frac{dx_N\,(t)}{dt} = f\,(x_N,t) + \eta_N(\omega,t), \tag{5.25}$$

where $\eta_N(\omega,t)$ *is the* N*-term approximation and* $\mathbb{E}\left[\int_0^T \eta_N(\omega,t)dt\right] < \infty$. *Then*

$$\lim_{N\to\infty} \mathbb{E}\left[|x\,(t) - x_N\,(t)|^2\right] \to 0, \tag{5.26}$$

uniformly in t, *iff* $x_N\,(t)$ *is the Karhunen Loève (KL) expansion of* $x\,(t)$.

Proof Throughout the proof, we will write $x(t)$ as $x(\omega,t)$ where $\omega \in \Omega$ is the sample space of the random process $x(t)$. First we will prove that if Eqn. (5.26) holds then $x_N\,(\omega,t)$ is the KL expansion of $x\,(\omega,t)$.

Let $\phi_m(t)$ be the set of any orthonormal basis, then $x\,(\omega,t)$ can be written as a convergent series as,

$$x\,(\omega,t) = \sum_{m=1}^{\infty} a_m(\omega)\phi_m(t),$$

without loss of generality we assume $a_m(\omega) = b_m c_m(\omega)$. Hence the above expansion becomes

$$x(\omega, t) = \sum_{m=1}^{\infty} b_m c_m(\omega)\phi_m(t). \tag{5.27}$$

Let the solution in Eqn. (5.25) be a $N-$term approximation of $x(\omega, t)$ which converge to true solution in mean square sense, given by

$$x_N(\omega, t) = \sum_{m=1}^{N} b_m c_m(\omega)\phi_m(t). \tag{5.28}$$

Hence the error e_N is given by

$$e_N(\omega, t) = \sum_{m=N+1}^{\infty} b_m c_m(\omega)\phi_m(t). \tag{5.29}$$

Projecting $x(\omega, t)$ in ϕ_m we get

$$c_m(\omega) = \frac{1}{b_m}\int_0^T x(\omega, t)\phi_m(t)dt. \tag{5.30}$$

Hence,

$$\begin{aligned}\mathbb{E}\left[e_N(\omega, t)\right] &= \sum_{m=N+1}^{\infty}\sum_{k=N+1}^{\infty}\phi_m(t)\phi_k(t)\int_0^T\int_0^T \mathbb{E}\left[x(\omega, t_1)\, x(\omega, t_2)\right]\phi_m(t_1)\phi_k(t_2)dt_1dt_2 \\ &= \sum_{m=N+1}^{\infty}\sum_{k=N+1}^{\infty}\phi_m(t)\phi_k(t)\int_0^T\int_0^T C_{xx}(t_1, t_2)\phi_m(t_1)\phi_k(t_2)dt_1dt_2, \end{aligned} \tag{5.31}$$

where $C_{xx}(t_1, t_2)$ is the covariance function of $x(\omega, t)$. For uniform convergence of error in t, the orthonormal basis $\phi_m(t)$ should minimize $\int_0^T \mathbb{E}\left[e_N(\omega, t)\right]dt$. So taking into account orthonormality of $\phi_m(t)$ we get

$$\int_0^T \mathbb{E}\left[e_N(\omega, t)\right]dt = \sum_{m=N+1}^{\infty}\int_0^T\int_0^T C_{xx}(t_1, t_2)\phi_m(t_1)\phi_m(t_2)dt_1dt_2, \tag{5.32}$$

which is to be minimized subject to

$$\int_0^T \phi_m(t)\phi_k(t)dt = \delta_{mk} \ \forall m, k \in \mathbb{N}. \tag{5.33}$$

So using the method of Lagrange multipliers our objective function is

$$\begin{aligned} F(\phi_m(t)) &= \min_{\phi_m(t)} \sum_{m=N+1}^{\infty} \int_0^T \int_0^T C_{xx}(t_1,t_2)\phi_m(t_1)\phi_m(t_2)dt_1dt_2 \\ &\quad -b_m^2\left(\int_0^T \phi_m(t)\phi_m(t)dt - 1\right), \end{aligned} \tag{5.34}$$

where b_m^2 are the Lagrange multipliers. Differentiating above equation with respect to $\phi_m(t)$ and setting the derivative to zero we get

$$\int_0^T \left[\int_0^T C_{xx}(t_1,t_2)\phi_m(t_1)dt_1 - b_m^2\phi_m(t_2)\right] dt_2 = 0 \tag{5.35}$$

which is satisfied when

$$\int_0^T C_{xx}(t_1,t_2)\phi_m(t_1)dt_1 = b_m^2\phi_m(t_2) \tag{5.36}$$

which is the Fredholm integral equation for the random process $x(\omega,t)$. Hence $\phi_m(t)$ and b_m^2 are the eigenfunctions and eigenvalues of $C_{xx}(t_1,t_2)$ respectively, which completes our proof.

Now we will prove that, if $x_N(\omega,t)$ is the KL expansion of $x(\omega,t)$, then to ensure that solution of Eqn. (5.25) converge to solution of Eqn. (5.24b) in mean square sense, $x_N(\omega,t)$ should be the solution of Eqn. (5.25). To prove this, we propose the following uniqueness conditions on solution of Eqn. (5.24b) and KL expansion of a random process.

Proposition 1 *Given the non-explosion condition and the Lipschitz condition are satisfied for $f(\cdot,\cdot)$ in Eqn. (5.24b). Let Z be a random variable, independent of*

the σ-algebra $\mathcal{F}_\infty^m$ generated by $\eta(\omega,t), t \geq 0$ and $\mathbb{E}\left[|Z|^2\right] \leq \infty$. Then the stochastic differential equation in Eqn. (5.24b) where $t \in [0,T], X(\omega,0) = Z$ has a unique t-continuous solution $x(\omega,t)$ with the property that $x(\omega,t)$ is adapted to the filtration $\mathcal{F}_t^Z$ generated by Z and $\eta(\omega,t), t \geq 0$ and $\mathbb{E}\left[\int_0^T |x(\omega,t)|^2 dt\right] \leq \infty$.

Proof See [116], Chapter 5.

Proposition 2 *The Karhunen Loève expansion of the random process $x(\omega,t)$ given by $x(\omega,t) = \sum_{m=1}^{\infty} \sqrt{\lambda_m}\xi_m(\omega)\phi_m(t)$ is unique.*

Proof See [25], Chapter 2.

Let us assume that $y_N(\omega,t)$ is the KL expansion of $x(\omega,t)$. Furthermore assume that $y_N(\omega,t) \neq x_N(\omega,t)$, which is the solution of Eqn. (5.25). Hence, we assume that although $y_N(\omega,t)$ is the KL expansion of $x(\omega,t)$ it does not satisfy Eqn. (5.25), whose solution converge to the solution of Eqn. (5.24b) in mean square sense.

Proposition 1 and 2 says that the solution of SDE is unique and any random process has an unique KL expansion. Also Eqn. (5.25) has unique solution as RHS of Eqn. (5.25) satisfies Lipschitz condition. This can be proven as follows: for right hand side of Eqn. (5.25) to be Lipschitz we must have

$$\begin{aligned} &|f(x,t) + \eta_N(\omega,t) - f(y,t) - \eta_N(\omega,t)| \leq C\,|x-y| \\ \Rightarrow &|f(x,t) - f(y,t)| \leq C\,|x-y|\,. \end{aligned}$$

which is true as $f(\cdot,\cdot)$ satisfies Lipschitz condition.

Hence Eqn. (5.24b) has unique solution which has an unique KL expansion. Also according to our assumption, the solution of Eqn. (5.25) converge to solution of Eqn. (5.24b) in mean square sense. This contradicts our assumption that $y_N(\omega,t) \neq x_N(\omega,t)$ as we have proved that Eqn. (5.25) has unique solutions, and for mean

square convergence in solutions of Eqn. (5.25) and Eqn. (5.24b), $x_N(\omega, t)$ must be the KL expansion of $x(\omega, t)$, which completes our proof.

Special Case Let $\eta_N(\omega, t)$ be the N-term KL expansion of $\eta(\omega, t)$. Given by,

$$\eta_N(\omega, t) = \sum_{m=1}^{N} \sqrt{\lambda_m^{\eta}} \xi_m^W(\omega) \phi_m^W(t). \tag{5.37}$$

Since $\eta_N(\omega, t)$ is the KL expansion, we have $\mathbb{E}\left[\int_0^T \eta_N(\omega, t) dt\right] < \infty$ [25]. Hence the solutions of Eqn. (5.25) converge to the solution of Eqn. (5.24b) in mean square sense. So, we suggest a corollary.

Corollary 2 *Given the stochastic differential equation in Eqn. (5.24b), where $f(\cdot, \cdot)$ satisfies the non-explosion and the Lipschitz condition, if the N-term approximation of $\eta(\omega, t)$ in Eqn. (5.25), $\eta_N(\omega, t)$, is given by the KL expansion of $\eta(\omega, t)$ then the solution of Eqn. (5.25) converges to the solution of Eqn. (5.24b) in mean square sense.*

Using corollary 2 we have showed that if the non-explosion and Lipschitz conditions are satisfied then the solution of Eqn. (5.25) converge to solution of Eqn. (5.24b) in mean square sense. In other words, if the process noise term in Eqn. (5.13) is approximated using KL expansion, then the solutions of Eqn. (5.15) and Eqn. (5.13) converge in mean square sense.

This proof gives us a framework to apply KL expansion to stochastic dynamical systems. In the next section, we will describe the methodology employed to predict evolution of probability density in stochastic dynamical systems using this framework.

D. Karhunen Loève Frobenius-Perron Formulation

Let us consider the dynamical system given in Eqn. (5.13) where $x \in \mathbb{R}^n$ are the states, having initial CDF $\pi(x,0) = \pi_0$. Let us also assume that the initial CDF admits a PDF given by ρ_0. The stochastic forcing term $\eta(\omega,t)$ in Eqn. (5.13) is approximated using KL expansion having finite number of terms (N). In the present case, where $\eta(\omega,t)$ is a Gaussian white noise, the approximation reduces to,

$$\eta\left(\omega,t\right) = \sqrt{2}\sum_{i=1}^{N}\xi_i\left(\omega\right)\cos\left(\left(i-\frac{1}{2}\right)\frac{\pi t}{T}\right), \tag{5.38}$$

After substituting the expression in Eqn. (5.38) in Eqn. (5.13) we get the differential Eqn. (5.15).

Using the methodology given in [129, 130] and as presented in the previous section, MOC is applied to Eqn. (4.7), and an augmented dynamical system is formed, with states $\left[x_N\left(t\right), \rho\left(x,t\right)\right]^T$, where $\rho\left(x,t\right)$ is the probability density of the states at time t, and $x_N\left(t\right) = \left[x_N^1\left(t\right), \ldots, x_N^n\left(t\right)\right]^T$. The augmented system is given by,

$$\dot{x}_N\left(t\right) = f\left(x_N,t\right) + \sqrt{2}\sum_{i=1}^{N}\xi_i\left(\omega\right)\cos\left(\left(i-\frac{1}{2}\right)\frac{\pi t}{T}\right), \tag{5.39a}$$

$$\dot{\rho}(x_N,t) = -\operatorname{div} f\left(x_N,t\right)\rho(x_N,t), \tag{5.39b}$$

where $\operatorname{div} f\left(x,t\right) = \sum_{i=1}^{n}\dfrac{\partial f\left(x,t\right)}{\partial x^i\left(t\right)}$. Equation (5.39a) & Eqn. (5.39b) can be solved to get the value of $\rho\left(x_N,t\right)$ along the characteristic curves of Eqn. (4.7). Detailed discussion of the solution methodology has been presented in the previous section.

1. Illustrative Examples

The proposed methodology is applied to a Vanderpol's oscillator, whose dynamics is governed by,

$$\ddot{x}(t) = \left(1 - x^2(t)\right)\dot{x}(t) - x(t) + \xi(t), \tag{5.40}$$

and to a Duffing oscillator, with dynamics,

$$\ddot{x}(t) = 10x(t) - 30x^3(t) - 10\dot{x}(t) + \xi(t), \tag{5.41}$$

with $\xi(t)$ having zero mean and autocorrelation $2\pi\mathbf{I}$. The initial state uncertainty, has a PDF $\rho_0(z) \sim \mathcal{N}\left([0,0]^T, \mathrm{diag}(1,1)\right)$ for both the systems in Eqn. (5.40) and Eqn. (5.41).

Let $x_1(t) = x(t)$ and $x_2(t) = \dot{x}(t)$, the augmented dynamical system for the Vanderpol's oscillator is given by,

$$\dot{x}_1(t) = x_2(t) \tag{5.42a}$$

$$\dot{x}_2(t) = \left(1 - x_1^2(t)\right)x_2(t) - x_1(t) + \xi(t) \tag{5.42b}$$

$$\dot{\rho}(x(t)) = -\left(1 - x_1^2(t)\right)\rho(x(t)), \tag{5.42c}$$

and for the Duffing oscillator, is given by,

$$\dot{x}_1(t) = x_2(t) \tag{5.43a}$$

$$\dot{x}_2(t) = 10x_1(t) - 30x_1^3(t) - 10x_2(t) + \xi(t) \tag{5.43b}$$

$$\dot{\rho}(x(t)) = 10\rho(x(t)). \tag{5.43c}$$

Next, the initial PDF, ρ_0 is sampled, with sample size of $M = 500$. For the Vanderpol's oscillator, final time T is 1s, for the Duffing oscillator $T = 3$s. Number of terms in the KL expansion is fixed to $N = 7$. Figure 54 shows the evolution of probability

densities with time for the two oscillators. The density value $\rho(x,t)$ is color coded with red representing high density value and blue representing low. It is observed that, for the Vanderpol's oscillator the probability mass accumulates along the limit cycle and for the Duffing oscillator, we get a bimodal PDF at final time. This is in agreement with the physical intuition of behavior of these systems.

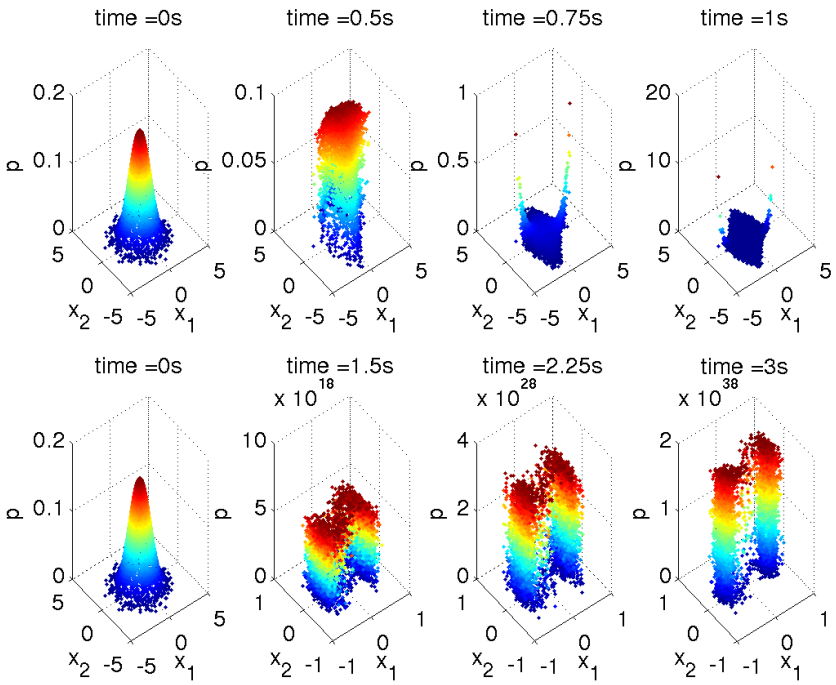
E. State Estimation Using Karhunen Loève Expansion and Frobenius-Perron Operator

We apply the proposed methodology of using Karhunen Loève and Frobenius-Perron (KLFP) operator, for uncertainty propagation to estimate states of a nonlinear system. The nonlinear estimation algorithm used is same as the one described in the previous chapter, that uses FP operator. First, we apply this to a Vanderpol's oscillator system, then we use the proposed methodology to estimate states of a hypersonic reentry vehicle. Here, we draw random particles (elements) from the domain $\mathbf{D}_0$ of $x(t=0)$. Propagation of uncertainty is done using Eqn. (5.39a) & Eqn. (5.39b). By using this framework for nonlinear estimation purposes, we get the exact value of prior PDF after propagation. This gives us significant advantage over particle filters where, generally the prior PDF is approximated using histograms. Bayesian inference is used to update the posterior PDF from prior for each particle.

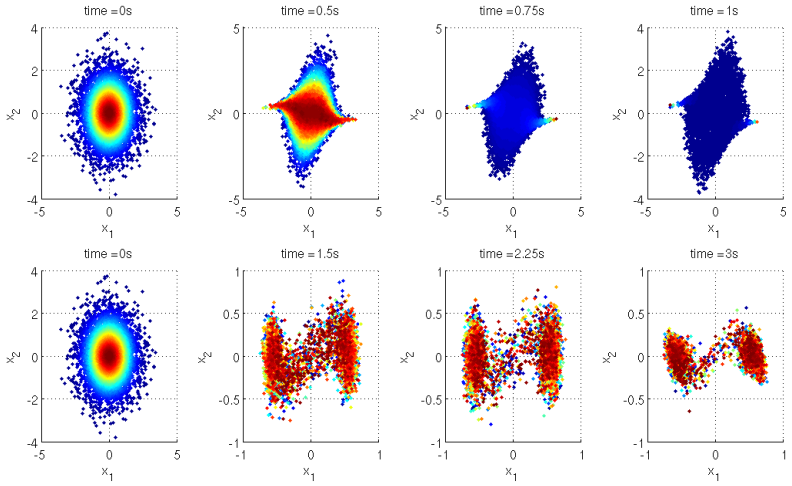
1. State Estimation of Vanderpol's Oscillator

In the present work, the system used is the Vanderpol's oscillator described by Eqn. (5.40). We use a nonlinear measurement model which is given by

$$y(t) = x_1^2(t) + x_2^2(t) + \zeta(t) \tag{5.44}$$



(a) Perspective view of the PDFs along direction $\left(\frac{\pi}{4}, \frac{\pi}{4}, \frac{\pi}{4}\right)$.



(b) Top view of the PDFs along direction $\left(\frac{\pi}{2}, \frac{\pi}{2}, 0\right)$.

Fig. 54.: Uncertainty propagation for Vanderpol's oscillator Eqn. (5.40) in top row of each figure, and Duffing oscillator Eqn. (5.41) in bottom row of each figure.

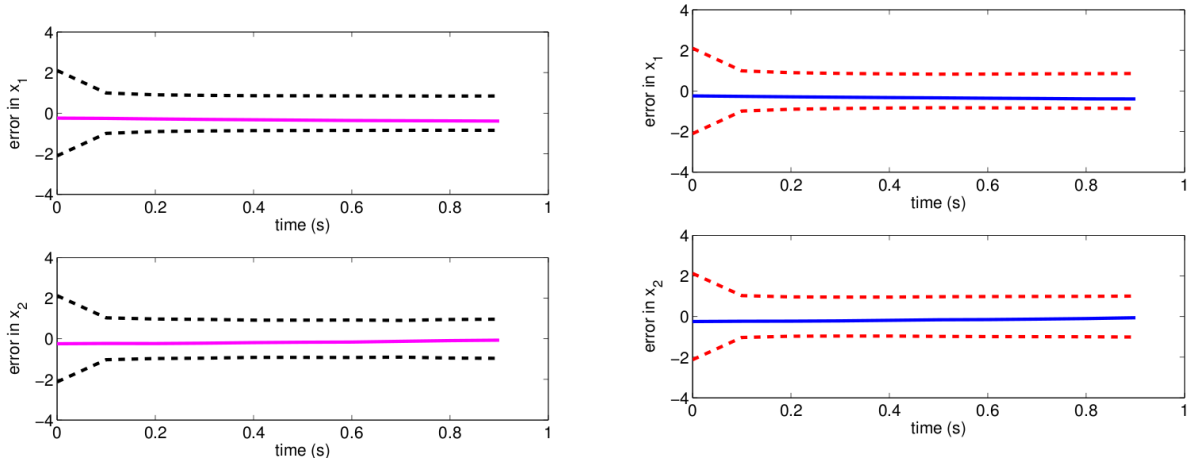
where $\zeta(t)$ is zero mean Gaussian measurement noise with autocorrelation $R = 6 \times 10^{-2}$. The process noise, $\eta(x,t)$ also considered to have zero mean and an autocorrelation of $Q = 6 \times 10^{-1}$.

We consider the initial state uncertainty to be normally distributed with $\mathbb{E}[x(0)] = [0,0]^T$ and $\mathbb{E}[x(0)^2] = \text{diag}(1,1)$. It is assumed that the initial states of the actual system is $[0.25, 0.25]$, with initial error in estimation being $[0.25, 0.25]$. The measurement update interval was fixed to 0.1s and final time T was assumed to be $T = 1$s.

The performance of the proposed filter is compared with the generic particle filter given in [9]. The sample size for each estimator is fixed to 500 elements. The simulations were performed on a Linux machine with Intel® Pentium D processor. Figure 55 shows the $\pm 3\sigma$ plots for the estimators. No major difference can be observed in the performance of the estimators as the errors in estimation are within the $\pm 3\sigma$ limits and converge. But if we compare the computational time, KLFP-based estimator takes 58.46s per filtering step, whereas the generic particle filter the time taken is 284.19s. Thus it can be concluded, given same sample size and the same processing environment, KLFP-based estimator achieves similar performance as the generic particle filter, and is almost five times faster than the particle filter based estimator.

2. Application to Hypersonic Reentry

The state estimation methodology is now applied to estimate states of a hypersonic reentry vehicle. The equations of motion are described by a six state Vinh's equation which is similar to the equations used in the previous section. Here we add process



(a) KLFP-based estimator. Magenta: Error in estimation, black: $\pm 3\sigma$ bounds.

(b) Generic particle filter. Blue- Error in estimation, red- $\pm 3\sigma$ bounds.

Fig. 55.: $\pm 3\sigma$ plots for a) KLFP-based estimator and b) Generic particle filter.

noise to states r and v. The equations with process noise are given by,

$$\dot{r} = v \sin\gamma + \eta_r, \tag{5.45a}$$

$$\dot{\theta} = \frac{v \cos\gamma \cos\xi}{r \cos\lambda}, \tag{5.45b}$$

$$\dot{\lambda} = \frac{v \cos\gamma \sin\xi}{r} \tag{5.45c}$$

$$\dot{v} = -\frac{\rho v^2}{2B_c} - g \sin\gamma - \Omega^2 r \cos\lambda(\sin\gamma \cos\lambda - \cos\gamma \sin\lambda \sin\xi) + \eta_v \tag{5.45d}$$

$$\dot{\gamma} = \left(\frac{v}{r} - \frac{g}{v}\right)\cos(\gamma) + \frac{\rho}{2B_c}\left(\frac{L}{D}\right) v \cos\sigma + 2\Omega \cos\lambda \cos\xi + \frac{\Omega^2 r}{v} \cos\lambda(\cos\gamma \cos\lambda + \sin\gamma \sin\lambda \sin\xi) \tag{5.45e}$$

$$\dot{\xi} = \frac{\rho}{2B_c}\left(\frac{L}{D}\right) v \sin\sigma - \frac{v}{r} \cos\gamma \cos\xi \tan\lambda + 2\Omega(\tan\gamma \cos\lambda \sin\xi - \sin\lambda) - \frac{\Omega^2 r}{v \cos\gamma} \sin\lambda \cos\lambda \cos\xi, \tag{5.45f}$$

where η_r and η_v are zero mean, Gaussian white noises with autocorrelation $Q = 6 \times 10^{-2}$, in scaled units. The other constants have the same values as in the previous

section. The initial condition uncertainty is assumed to be Gaussian with mean and covariance given by,

$$\mu_0 = [R_m + 54 \text{ Km}, -60^o, 30^o, 2.4 \text{ Km/s}, -9^o, 0.0573^o]^T \tag{5.46a}$$

$$\sigma_0 = \begin{bmatrix} 5.4 \text{ Km} & 0 & 0 & 0 & 0 & 0 \\ 0 & 3^o & 0 & 0 & 0 & 0 \\ 0 & 0 & 3^o & 0 & 0 & 0 \\ 0 & 0 & 0 & 240 \text{ m/s} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.9^o & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0057^o \end{bmatrix}. \tag{5.46b}$$

A nondimensionalized system is constructed by scaling each term in Eqn. (5.45a) through Eqn. (5.45f) with scaling constants. We use the same scaling constants as used in previous two section which are given in Table IV.

The measurement model consists of the dynamic pressure $\bar{q}$ the heating rate H, the flight path angle γ, and measurements along geocentric longitude and latitude θ and λ. Here we add another observation where we observe the azimuth angle ξ. Hence

$$\tilde{y} = \begin{pmatrix} \bar{q} & H & \gamma & \theta & \lambda & \xi \end{pmatrix} \tag{5.47}$$

The constants scaling the measurements are given in Table VIII. We assume measurement noise added to each state with constant autocorrelation $R = 6 \times 10^{-1}$.

The true states are assumed to be $[R_m + 61 \text{ km}, -60.6^o, 30.3^o, 2.42 \text{ km/s}, -9.09^o, 0.0573^o]$, hence the estimation algorithms start with an initial error. The measurement update interval is kept fixed at 10 seconds.

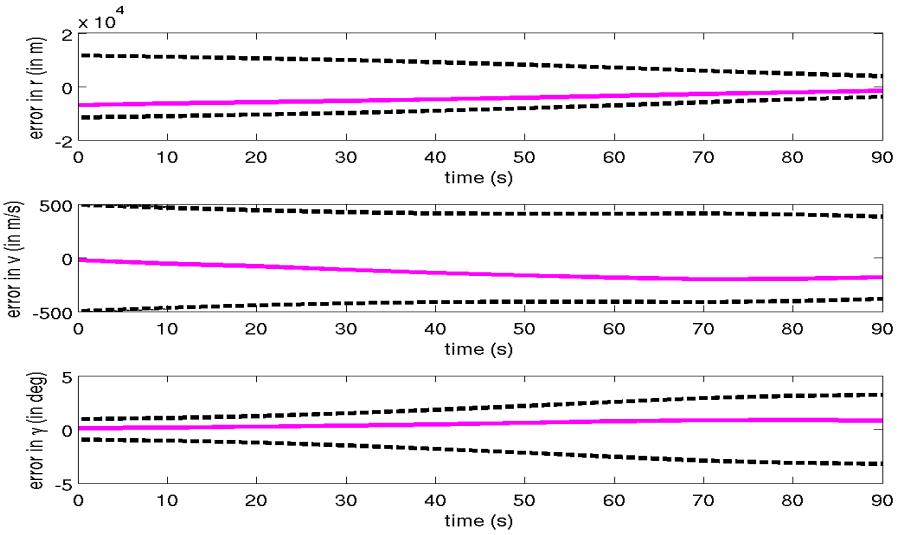
Figure 56 shows plots for $\pm 3\sigma$ limits and estimation error for particle filter. We can see that the errors diverge and the $\pm 3\sigma$ limits do not converge, for all the states

Table VIII.: Normalization factors used for measurements (six state Vinh's equation with noise)

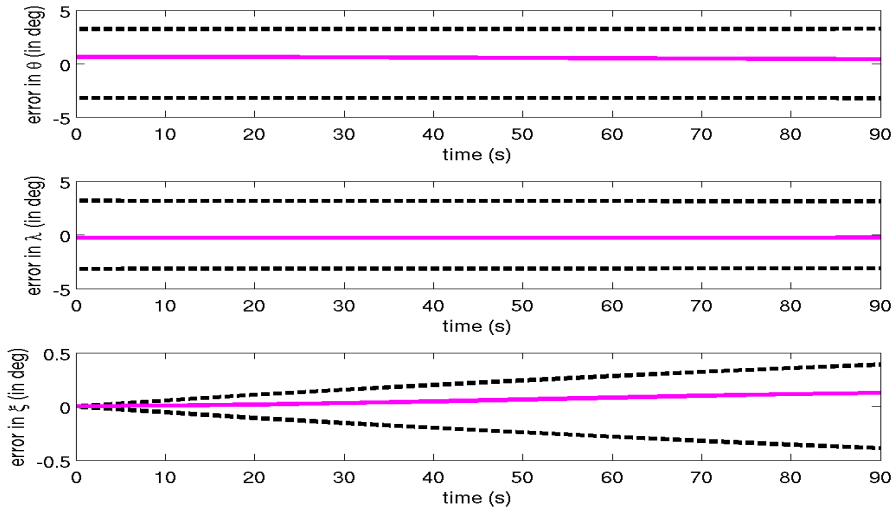
Measurement	**Normalization Factors**
Dynamic Pressure	$1.97 \times 10^3 \frac{\text{N}}{\text{m}}$
Heating Rate	$0.0231 \frac{\text{J}}{\text{m-s}}$
Flight Path Angle	9.6422^o
Geocentric Longitude	60.6^o
Geocentric Latitude	30.3^o
Azimuth Angle	1.9966^o

except for r. Figure 57 uses the proposed estimation methodology with KL expansion and FP operator, it can be observed that the proposed method performs better than the particle filter in minimizing the covariance and reducing the estimation errors. Hence it can be concluded that for the particular application, the proposed estimator performs better than particle filter. However, it can be observed that variance of ξ for both the estimator increase, and though the proposed estimator is better than the particle filter in minimizing variance its performance can not be said to be optimal.

As mentioned in Chapter II the $\pm 3\sigma$ do not reveal much regarding the efficacy of a filtering algorithm when we have a nonlinear system with non-Gaussian PDF evolution. Hence we plot the univariate and bivariate posterior marginals obtained after state estimation by the two estimation methods. For particle filter, we use histogram technique to calculate the joint PDF and then integrate out all other states except for the state for which marginal is to be calculated [131]. For the proposed estimator using KL expansion and FP operator, we first divide the domain of the random of the random variable into equally sized bins. Then we use the average of

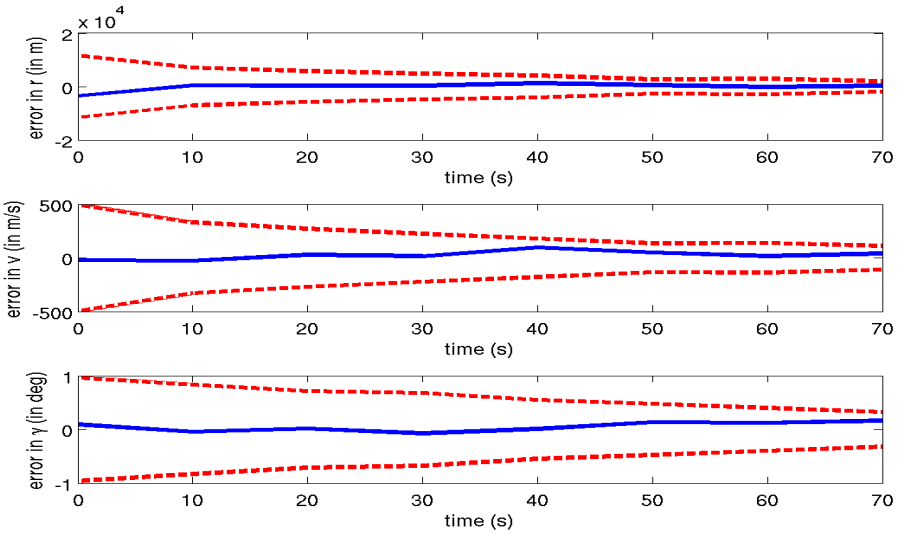


(a) Plots for r, v and γ.

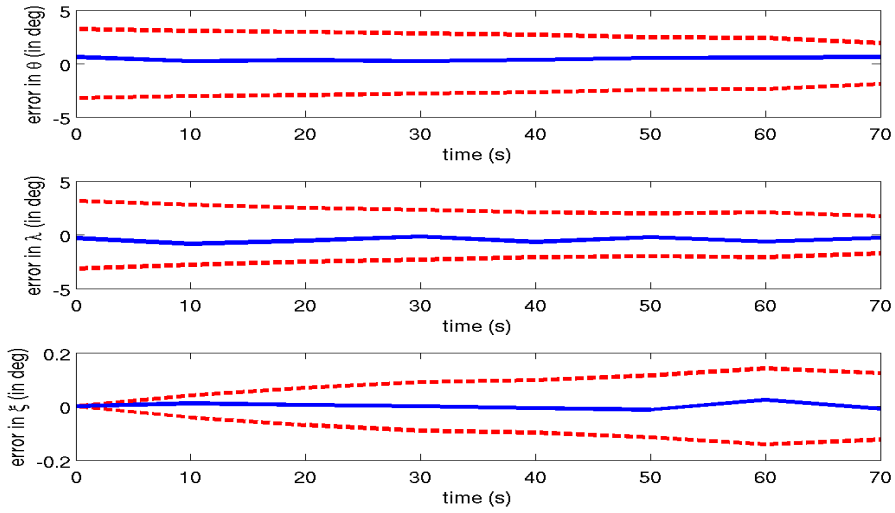


(b) Plots for θ, λ and ξ.

Fig. 56.: Estimation errors (solid lines) and $\pm 3\sigma$ limits (dashed lines) for particle filter based estimator.



(a) Plots for r, v and γ.



(b) Plots for θ, λ and ξ.

Fig. 57.: Estimation errors (solid lines) and $\pm 3\sigma$ limits (dashed lines) for estimator using KL expansion and FP operator.

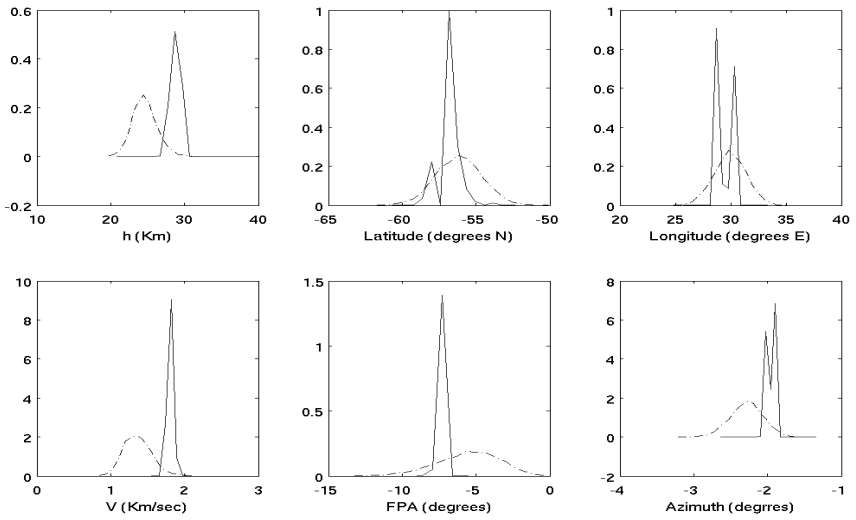


Fig. 58.: Plot for univariate marginal density for particle filter (dashed line) and KL expansion and FP operator based estimator (solid line). Here $h = r - R_m$ and the y-axis denotes PDF value.

the density values of points, obtained after using FP operator, in a bin, to get the PDF value for that particular bin. Detailed description of this method can be found in ref. [129].

Figure 58 and Fig. 59 show plots for univariate and bivariate marginal density respectively, for the two estimators compared in this chapter. There can be 15 state combinations for bivariate densities, of which we show four combinations. It can be observed that the dispersion of the particle filter based estimator is more than the proposed estimator as the histogram approximation of PDF tend to smear out the probability mass, whereas the proposed estimator captures the concentration of the probability mass well. Hence it can be seen that proposed estimator is more successful in capturing the actual reduction uncertainty than the particle filter.

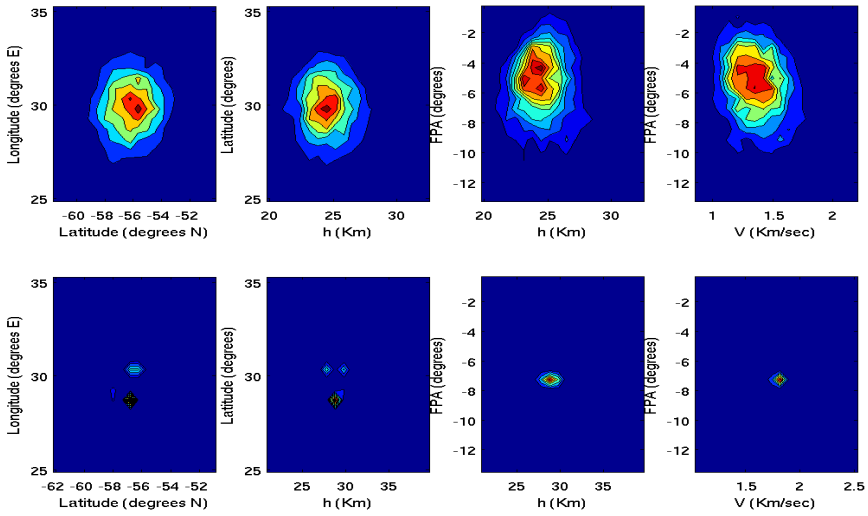


Fig. 59.: Plot for bivariate marginal density for particle filter (top row) and KL expansion and FP operator based estimator (bottom row) for 4 state combinations. Here $h = r - R_m$. The PDF value is color coded, red is high PDF and blue is low.

F. Summary

In this section, we present an uncertainty quantification methodology, where the process noise term is expanded using KL expansion and uncertainty is propagated using FP operator. We have shown that if in a stochastic dynamical system if the noise term is approximated using finite term KL expansion the solution of the approximated dynamics converge to the actual solution in mean square sense. Moreover, we propose an estimation technique based on Kl expansion and FP operator, and show that the proposed estimator outperforms particle filter in terms of performance.

CHAPTER VI

CONCLUSIONS AND SCOPE OF FUTURE WORK

This dissertation has presented novel methods of uncertainty quantification of nonlinear system with emphasis on state estimation problems. Most of the real-world systems deal with the problems of high dimensionality, nonlinearity and high uncertainty. Hence, characterization of uncertainty is a difficult task in practice for these systems. One of the major problems in this scenario is estimation of states and parameters. Most commonly used estimation algorithms, involve local linearization of dynamics and assume Gaussian PDF evolution. But due to high nonlinearity of the real-world systems, estimation becomes a difficult task to achieve.

There are estimation algorithms, based on Monte Carlo methods which alleviate the problem of nonlinearity. But they have been observed to fail if the state space dimension is high.

In this work we use the uncertainty quantification methods developed to estimate states and parameters of a nonlinear system. We first present two methods based on polynomial chaos framework. Polynomial chaos is a method where the any random process is expanded as a linear combination of orthogonal polynomials of the underlying random variable. The estimation algorithms use polynomial chaos for forward propagation of uncertainty, and the update step or the inverse problem is solved using higher order moment updates, and Bayesian inference. The proposed method was applied to estimate states of a duffing oscillator and vehicle reentering the atmosphere of Mars, and its performance was compared with EKF based estimators. It was found that the performance of the PC based estimators was superior than EKF for the particular application.

Next we presented an algorithm which uses Frobenius-Perron (FP) operator the-

ory for propagation of uncertainty. The FP operator given by a partial differential equation, which dictates the evolution of densities in a dynamical system with parametric uncertainty. The power of FP operator is that we know the exact density values at each sample points after propagation. We have formulated an estimation algorithm based on FP operator and compared the results to particle filters. It was found that the FP operator is computationally superior than the particle filtering based estimators.

Finally we have proposed an estimation algorithm, when the dynamical system has process noise in it. We use the Karhunen Loève (KL) expansion to represent process noise and then use FP operator based methodology to propagate uncertainty. We have proved that the solutions of the stochastic differential equation when the noise is approximated using KL expansion converge in mean square sense to the actual solutions. Finally we have applied the proposed estimation algorithm to estimate states of a duffing oscillator and Vinh's equation. It was found that the proposed estimation method is superior to the particle filtering method for the particular application.

A. Future Work

There are several directions in which one can work towards, from the methodologies and algorithms developed in this dissertation. In this section we will discuss in brief some of the future directions of research.

1. Stochastic Control

One of the main use of the methodologies of uncertainty quantification developed in this dissertation is control of stochastic dynamical system. Up until now, stochastic control methodologies mainly focus on minimizing the expectation of the underlying

cost function, amongst them the linear quadratic Gaussian formulation is widely used [132]. But there are several systems for which it is difficult to minimize a cost function based on expected value. In such cases, one can resort to providing performance guarantees in a weaker sense like in distribution or probabilistic guarantees. The FP operator technique can play an important role in such cases. One may use the fact that the uncertainty propagation is exact in this case.

2. Multiphysical Dynamical Systems

More often than not we are faced with systems which are multiphysical in nature. Dynamics in such systems evolve in different scales and are spatially and temporally varying. The dynamical system in such cases is described by a partial differential equation (PDE) [133]. Uncertainty management in such systems in a difficult problems and faces a lot of challenges, as the measurements are available in different scales and of different quantities than which we want to estimate. Most of the methods for quantification of uncertainty involve Monte Carlo (MC) simulations and suffer from the curse of dimensionality [134]. We have observed that FP operator based method is computationally superior than MC methods and predicts the uncertainty accurately. The use of FP operator for parameter estimation and sensitivity analysis of dynamical systems may help reduce the computational burden due to use Monte Carlo and PDF approximation using histograms.

3. Dimensional Scaling

Almost all of the uncertainty quantification (UQ) algorithms suffer from the curse of dimensionality. For MC methods the number of sample points needed to obtain a sufficiently good estimate of the quantity of interest increase exponentially with increase in dimension. Though FP operator was found to better than MC methods for UQ

in the particular application discussed in this dissertation, it fails if the dimensionality of the state space is very high. We can use dimensional scaling or dimensional interpolation in such cases.

Dimensional scaling generalizes any system into a N dimensional representation, and has been widely used in chemical physics [135]. It has been successfully applied to solve the Schrödinger's equation to find the ground state energy of a hydrogen atom [136]. A technique has been intuitively proposed by Daum [137], where dimensional scaling is used alongside FP operator for UQ in high dimensional systems. We can use this technique for parameter estimation and uncertainty quantification of several nonlinear systems, with large state space and high uncertainty.

REFERENCES

[1] R.E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[2] B. Anderson, J.B. Moore, and M. Eslami, "Optimal filtering," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 12, no. 2, pp. 235–236, 1982.

[3] G. Evensen, "Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model," *Journal of Geophysical Research*, vol. 97, no. C11, pp. 17905–17924, 1992.

[4] R.N. Miller, M. Ghil, and F. Gauthiez, "Advanced data assimilation in strongly nonlinear dynamical systems," *Journal of the Atmospheric Sciences*, vol. 51, no. 8, pp. 1037–1056, 1994.

[5] P. Gauthier, P. Courtier, and P. Moll, "Assimilation of simulated wind lidar data with a Kalman filter," *Monthly Weather Review*, vol. 121, no. 6, pp. 1803–1820, 1993.

[6] F. Bouttier, "A dynamical estimation of forecast error covariances in an assimilation system," *Monthly Weather Review*, vol. 122, no. 10, pp. 2376–2390, 1994.

[7] S.J. Julier and J.K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Proceedings of Int. Symp. Aerospace/Defense Sensing, Simul. and Controls*, 1997, vol. 3, pp. 182–193.

[8] N.J. Gordon, D.J. Salmond, and A.F.M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *IEE Proceedings on Radar and Signal Processing*, 1993, vol. 140, pp. 107–113.

[9] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[10] A. Khalil, M.and Sarkar and S. Adhikari, "Nonlinear filters for chaotic oscillatory systems," *Nonlinear Dynamics*, vol. 55, pp. 113–137, 2009.

[11] P.W. Glynn and D.L. Iglehart, "Importance sampling for stochastic simulations," *Management Science*, pp. 1367–1392, 1989.

[12] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 1–25, 1996.

[13] A. Doucet, S.J. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

[14] H. Tanizaki, *Nonlinear Filters: Estimation and Applications*, Berlin: Springer-Verlag, 1996.

[15] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Norwood, MA: Artech House Publishers, 2004.

[16] A. Doucet, N. De Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*, New York: Springer-Verlag, 2001.

[17] F. Daum, J. Huang, and R. Co, "Curse of dimensionality and particle filters," in *Proceedings of IEEE Aerospace Conference*, 2003, pp. 4_1979–4_1993.

[18] C.S. Manohar and D. Roy, "Monte Carlo filters for identification of nonlinear structural dynamical systems," *Sadhana Acad. Proc. Eng. Sciences*, vol. 31, pp. 399–427, 2006.

[19] J.S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *Journal of the American Statistical Association*, vol. 93, no. 443, pp. 1032–1044, 1998.

[20] N. Oudjane and C. Musso, "Progressive correction for regularized particle filters," in *Proceedings of the Third International Conference on Information Fusion*. IEEE, 2000, vol. 2, pp. THB2–10.

[21] W.R. Gilks and C. Berzuini, "Following a moving target Monte Carlo inference for dynamic Bayesian models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 1, pp. 127–146, 2001.

[22] C. Snyder, T. Bengtsson, P. Bickel, and J. Anderson, "Obstacles to high-dimensional particle filtering," *Monthly Weather Review*, vol. 136, no. 12, pp. 4629–4640, 2008.

[23] G. Casella and C.P. Robert, "Rao-Blackwellisation of sampling schemes," *Biometrika*, vol. 83, no. 1, pp. 81–94, 1996.

[24] F. Daum and R. Co, "Nonlinear filters: beyond the Kalman filter," *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, no. 8 Part 2, pp. 57–69, 2005.

[25] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, New York: Springer-Verlag, 1991.

[26] A. Lasota and M.C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, New York: Springer-Verlag, 1994.

[27] N. Wiener, "The homogeneous chaos," *American Journal of Mathematics*, vol. 60, no. 4, pp. 897–936, 1938.

[28] R. H. Cameron and W. T. Martin, "The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals," *The Annals of Mathematics*, vol. 48, no. 2, pp. 385–392, 1947.

[29] D. Xiu and G. E. Karniadakis, "The Wiener–Askey polynomial chaos for stochastic differential equations," *SIAM J. Sci. Comput.*, vol. 24, no. 2, pp. 619–644, 2002.

[30] R. Askey and J. Wilson, "Some basic hypergeometric polynomials that generalize Jacobi polynomials," *Memoirs Amer. Math. Soc.*, vol. 54, no. 314, pp. 1–4, 1985.

[31] T. Y. Hou, W. Luo, B. Rozovskii, and H-M. Zhou, "Wiener chaos expansions and numerical solutions of randomly forced equations of fluid mechanics," *J. Comput. Phys.*, vol. 216, no. 2, pp. 687–706, 2006.

[32] D. Xiu and G. E. Karniadakis, "Modeling uncertainty in flow simulations via generalized polynomial chaos," *J. Comput. Phys.*, vol. 187, no. 1, pp. 137–167, 2003.

[33] R. Ghanem and J. Red-Horse, "Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach," *Phys. D*, vol. 133, no. 1-4, pp. 137–144, 1999.

[34] R.G. Ghanem, "Ingredients for a general purpose stochastic finite elements implementation," *Comput. Methods Appl. Mech. Eng.*, vol. 168, no. 1-4, pp. 19–34, 1999.

[35] C. Sandu E. D. Blanchard, A. Sandhu, "A polynomial chaos based Bayesian approach for estimating uncertain parameters of mechanical systems part 1: Theoretical approach," Tech. Rep. TR-07-38, Computer Science Department, Virginia Polytechnic Inst. and State Univ., Blacksburg, VA, 2007.

[36] C. Sandu E. D. Blanchard, A. Sandhu, "A polynomial chaos based Bayesian approach for estimating uncertain parameters of mechanical systems part 2: Application to vehicle systems," Tech. Rep. TR-07-39, Computer Science Department, Virginia Polytechnic Inst. and State Univ., Blacksburg, VA, 2007.

[37] R. Kubo, "Stochastic Liouville equations," *Journal of Mathematical Physics*, vol. 4, pp. 174–183, 1963.

[38] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, Berlin: Springer-Verlag, 1996.

[39] T. Runolfsson and C. Lin, "Computation of uncertainty distributions in complex dynamical systems," in *Proceedings of the American Control Conference*. IEEE Press, 2009, pp. 2458–2463.

[40] F. Daum and M. Krichman, "Non-particle filters," in *Proceedings of SPIE, the International Society for Optical Engineering*, 2006, vol. 6236, pp. 623614_1–623614_12.

[41] J. Kuczka, P. Hanggi, and A. Gadomski, "Non-Markovian process driven by quadratic noise: Kramers-Moyal expansion and Fokker-Planck modeling,"

Physical Review-Section E-Statistical Physics Plasma Fluids Related Interdiscpl Topics, vol. 51, no. 4, pp. 2933–2938, 1995.

[42] AN Kolmogorov, "On analytical methods in the theory of probability," *Uspehi. Mat. Nauk*, vol. 1938, pp. 5–41, 1938.

[43] M.F. Wehner and W.G. Wolfer, "Numerical evaluation of path-integral solutions to Fokker-Planck equations," *Physical Review A*, vol. 27, no. 5, pp. 26632670, 1983.

[44] HP Langtangen, "A general numerical solution method for Fokker-Planck equations with applications to structural reliability," *Probabilistic Engineering Mechanics*, vol. 6, no. 1, pp. 33–48, 1991.

[45] M. Di Paola and A. Sofi, "Approximate solution of the Fokker-Planck-Kolmogorov equation," *Probabilistic Engineering Mechanics*, vol. 17, no. 4, pp. 369–384, 2002.

[46] SF Wojtkiewicz and LA Bergman, "Numerical solution of high dimensional Fokker-Planck equations," in *8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability, South Bend, IN*, 2000, pp. PMC2000–167_1–6.

[47] GW Wei, "A unified approach for the solution of the Fokker-Planck equation," *Journal of Physics A: Mathematical and General*, vol. 33, pp. 4935–4953, 2000.

[48] A. Masud and L.A. Bergman, "Application of multi-scale finite element methods to the solution of the Fokker-Planck equation," *Computer Methods in Applied Mechanics and Engineering*, vol. 194, no. 12-16, pp. 1513–1526, 2005.

[49] M. Kumar, P. Singla, S. Chakravorty, and J.L. Junkins, "The partition of unity finite element approach to the stationary Fokker-Planck equation," in *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Keystone, CO, August 2006, AIAA, Paper 2006-6285.

[50] S.L. Shapiro, "Monte Carlo simulations of the 2+ 1 dimensional Fokker-Planck equation-Spherical star clusters containing massive, central black holes," in *Dynamics of Star Clusters: Proceedings of the 113th Symposium of the International Astronomical Union*, Princeton, NJ, 1985, vol. 113, pp. 373–412.

[51] K. Kikuchi, M. Yoshida, T. Maekawa, and H. Watanabe, "Metropolis Monte Carlo method as a numerical technique to solve the Fokker–Planck equation," *Chemical Physics Letters*, vol. 185, no. 3-4, pp. 335–338, 1991.

[52] D.W. Scott, *Multivariate Density Estimation*, vol. 139, New York: Wiley Online Library, 1992.

[53] M. Magdziarz and A. Weron, "Fractional Fokker-Planck dynamics: Stochastic representation and computer simulation," *Phys Rev. E*, vol. 75, no. 1 Pt 2, pp. 016708_1–016708_6, 2007.

[54] H. Singer, "Parameter estimation of nonlinear stochastic differential equations: simulated maximum likelihood versus extended Kalman filter and Itô-Taylor expansion," *Journal of Computational and Graphical Statistics*, vol. 11, no. 4, pp. 972–995, 2002.

[55] M. Kumar, S. Chakravorty, and J.L. Junkins, "A semianalytic meshless approach to the Fokker-Planck equation with application to hybrid systems," in *IEEE Conference on Decision and Control*. IEEE, 2007, pp. 3078–3083.

[56] J. Elf, P. Lötstedt, and P. Sjöberg, "Problems of high dimension in molecular biology," in *Proceedings of the 19th GAMM-Seminar, Leipzig, Germany*, 2003, pp. 21–30.

[57] J. Rust, "Using randomization to break the curse of dimensionality," *Econometrica*, vol. 65, no. 3, pp. 487–516, 1997.

[58] H.J. Bungartz and M. Griem, "Sparse grids," *Acta Numerica*, vol. 13, no. 1, pp. 147–269, 2004.

[59] J. Garcke and M. Griebel, "On the computation of the eigenproblems of hydrogen and helium in strong magnetic and electric fields with the sparse grid combination technique," *Journal of Computational Physics*, vol. 165, no. 2, pp. 694–716, 2000.

[60] R. Bellman, *Dynamic Programming*, Princeton, NJ: Princeton Univ. Press, 1957.

[61] K. Karhunen, "Über lineare methoden in der wahrscheinlichkeitsrechnung," *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.*, , no. 37, pp. 1–79, 1947.

[62] M. Loève, *Probability Theory, Vol. II*, New York: Springer-Verlag, 1978.

[63] G. Adomian, *Nonlinear Stochastic Systems Theory and Applications to Physics*, Norwood, MA: Kluwer Academic Pub, 1989.

[64] D.H. Chambers, R.J. Adrian, P. Moin, D.S. Stewart, and H.J. Sung, "Karhunen–Loève expansion of Burger's model of turbulence," *Physics of Fluids*, vol. 31, pp. 2573–2582, 1988.

[65] A.J. Newman, "Model reduction via the Karhunen-Loève expansion Part 1: An exposition," Tech. Rep., Institute for Systems Research, University of Maryland, College Park, MD, 1996.

[66] A.J. Newman, "Model reduction via the Karhunen-Loève expansion Part 2: Some elementary examples," Tech. Rep., Institute for Systems Research, University of Maryland, College Park, MD, 1996.

[67] S. Lall, J.E. Marsden, and S. Glavaski, "A subspace approach to balanced truncation for model reduction of nonlinear control systems," *International Journal of Robust and Nonlinear Control*, vol. 12, no. 6, pp. 519–535, 2002.

[68] R. Ghanem and P.D. Spanos, "A stochastic Galerkin expansion for nonlinear random vibration analysis," *Probabilistic Engineering Mechanics*, vol. 8, no. 3-4, pp. 255–264, 1993.

[69] G.W. Wornell, "A Karhunen-Loève-like expansion for 1/f processes via wavelets," *IEEE Transactions on Information Theory*, vol. 36, no. 4, pp. 859–861, 1990.

[70] K.K. Phoon, S.P. Huang, and S.T. Quek, "Implementation of Karhunen-Loève expansion for simulation using a wavelet-Galerkin scheme," *Probabilistic Engineering Mechanics*, vol. 17, no. 3, pp. 293–303, 2002.

[71] K.W. Yip and T.S. Ng, "Karhunen-Loève expansion of the WSSUS channel output and its application to efficient simulation," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 4, pp. 640–646, 1997.

[72] C.F. Gauss, *Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections*, Boston: Little, Brown and Comp., 1857.

[73] J.L. Crassidis and J.L. Junkins, *Optimal Estimation of Dynamic Systems*, vol. 2, Boca Raton, FL: Chapman & Hall, 2004.

[74] C.R. Rao, "Minimum variance quadratic unbiased estimation of variance components," *Journal of Multivariate Analysis*, vol. 1, no. 4, pp. 445–456, 1971.

[75] H. W. Sorenson, "Least-squares estimation: from Gauss to Kalman," *IEEE Spectrum*, vol. 7, pp. 63–68, July 1970.

[76] R.A. Fisher, "On an absolute criterion for fitting frequency curves," *Statistical Science*, vol. 12, no. 1, pp. 39–41, 1997.

[77] N. Wiener, "Extrapolation, interpolation, and smoothing of stationary time series with engineering applications," *Bull. Amer. Math. Soc. 56 (1950), 378-381*, vol. 2, no. 9904, pp. 09416–6, 1950.

[78] M. Bayes and M. Price, "An essay towards solving a problem in the doctrine of chances," *Royal Society of London Philosophical Transactions Series I*, vol. 53, pp. 370–418, 1763.

[79] R.E. Kalman and R.S. Bucy, "New results in linear filtering and prediction theory," *Transactions of the ASME. Series D, Journal of Basic Engineering*, vol. 83, pp. 95–107, 1961.

[80] C.E. Rasmussen, "The infinite Gaussian mixture model," *Advances in Neural Information Processing Systems*, vol. 12, pp. 554–560, 2000.

[81] T.W. Lee and M.S. Lewicki, "The generalized Gaussian mixture model using ICA," in *International Workshop on Independent Component Analysis*, Helsinki, Finland, 2000, pp. 239–244.

[82] R. Chen and J.S. Liu, "Mixture Kalman filters," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 62, no. 3, pp. 493–508, 2000.

[83] P. Dutta and R. Bhattacharya, "Nonlinear estimation with polynomial chaos and higher order moment updates," in *Proceedings of American Control Conference*, Baltimore, MD, 2010, IEEE, pp. 3142–3147.

[84] E.A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proceedings of The IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium*, Lake Louise, AB, Canada, 2000, IEEE, pp. 153–158.

[85] J.S. Liu, R. Chen, and T. Logvinenko, "A theoretical framework for sequential importance sampling and resampling," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. De Freitas, and N. Gordon, Eds., pp. 225–246. Springer-Verlag, New York, 2001.

[86] G.B. Arfken and H.J. Weber, *Mathematical Methods for Physicists*, New York: Academic Press, 2001.

[87] B. J. Debusschere, H. N. Najm, P. P. Pébay, O. M. Knio, R. G. Ghanem, and O. P. Le Maître, "Numerical challenges in the use of polynomial chaos representations for stochastic processes," *SIAM J. Sci. Comput.*, vol. 26, no. 2, pp. 698–719, 2005.

[88] J. Fisher and R. Bhattacharya, "Linear quadratic regulation of systems with stochastic parameter uncertainties," *Automatica*, vol. 35, no. 12, pp. 2831–2841, December 2009.

[89] S. Julier, J. Uhlmann, and H.F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, 2000.

[90] S.J. Julier and J.K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[91] R.S. Park and D.J. Scheeres, "Nonlinear mapping of Gaussian statistics: theory and applications to spacecraft trajectory design," *Journal of Guidance Control and Dynamics*, vol. 29, no. 6, pp. 1367–1375, 2006.

[92] M. Majji, J. Junkins, and J. D. Turner, " J^{th} moment extended Kalman filtering for estimation of nonlinear dynamic systems," in *AIAA Guidance, Navigation and Control Conference and Exhibit*, Honolulu, HI, August 2008, AIAA, Paper 2008-7386.

[93] G. Casella, R.L. Berger, and R.L. Berger, *Statistical Inference*, Grove, CA: Duxbury Pacific, 2002.

[94] B.D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge: Cambridge University Press, 2008.

[95] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832–837, 1956.

[96] N. Bourbaki, *Elements of Mathematics: Commutative Algebra*, New York: Springer Verlag, 1998.

[97] A. E. Bryson and Y. C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control*, New York: Taylor & Francis, 1975.

[98] P.E. Gill, W. Murray, and M.A. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization," *SIAM Journal on Optimization*, vol. 12, no. 4, pp. 979–1006, 2002.

[99] J.S. Chern and N.X. Vinh, "Optimal reentry trajectories of a lifting vehicle," Tech. Rep., NASA-CR-3236, NASA, 1980.

[100] P. Sengupta and R. Bhattacharya, "Uncertainty analysis of hypersonic flight using multi-resolution Markov operators," in *AIAA Guidance Navigation and Control Conference and Exhibit*, Honolulu, HI, August 2008, AIAA, Paper 2008-6298.

[101] K.P. Bollino, I.M. Ross, and D.D. Doman, "Optimal nonlinear feedback guidance for reentry vehicles," in *AIAA Guidance Navigation and Control Conference and Exhibit*, Keystone, CO, August 2006, AIAA, Paper 2006-6074.

[102] A. Prabhakar, J. Fisher, and R. Bhattacharya, "Polynomial chaos based analysis of probabilistic uncertainty in hypersonic flight dynamics," *Journal of Guidance, Control, and Dynamics*, vol. 33, no. 1, pp. 222–234, 2010.

[103] A.D. Polyanin and V.F. Zaitsev, *Handbook of Nonlinear Partial Differential Equations*, Boca Raton, FL: CRC Press, 2004.

[104] R. Franke, "Scattered data interpolation: Tests of some methods," *Math. Comput.*, vol. 38, no. 157, pp. 181–200, 1982.

[105] J.H. Halton, "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals," *Numerische Mathematik*, vol. 2, no. 1, pp. 84–90, 1960.

[106] F. Nobile, R. Tempone, and C.G. Webster, "A sparse grid stochastic collocation method for partial differential equations with random input data," *SIAM J. Numer. Anal.*, vol. 46, no. 5, pp. 2309–2345, 2008.

[107] C. Zenger, "Sparse grids," Tech. Rep. TUM-I–9037, Inst. fuer Informatik, Muenchen Technische Univ., Germany, Munich, Germany, 1997, `http://hdl.handle.net/10068/146730`.

[108] R. Becker and R. Rannacher, "An optimal control approach to *a posteriori* error estimation in finite element methods," *Acta Numerica*, vol. 10, pp. 1–102, 2003.

[109] M.B. Giles and E. Süli, "Adjoint methods for pdes: *a posteriori* error analysis and post-processing by duality," *Acta Numerica*, vol. 11, pp. 145–236, 2003.

[110] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Basel, Switzerland: Birkhäuser Verlag, 2003.

[111] F. Daum, M. Krichman, and R. Co, "Meshfree adjoint methods for nonlinear filtering," in *Proc. of IEEE Aerospace Conference*, Woburn,MA, July 2006, pp. 1–16, doi:10.1109/AERO.2006.1655925.

[112] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Boca Raton, FL: Chapman and Hall, 1996.

[113] C.W. Gardiner, *Handbook of Stochastic Methods*, Berlin: Springer Verlag, 1985.

[114] E. Pardoux and S.G. Peng, "Adapted solution of a backward stochastic differential equation," *Systems & Control Letters*, vol. 14, no. 1, pp. 55–61, 1990.

[115] P.E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Berlin: Springer-Verlag, 1992.

[116] B.K. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, Berlin: Springer-Verlag, 2003.

[117] K. Itô and M. Nisio, "On stationary solutions of a stochastic differential equation," *Kyoto Journal of Mathematics*, vol. 4, no. 1, pp. 1–75, 1964.

[118] N. Wiener, *Nonlinear Problems in Random Theory*, Cambridge, MA: Technology Press of Massachusetts Institute of Technology, 1958.

[119] F. Liu, V. Anh, and I. Turner, "Numerical solution of the space fractional Fokker-Planck equation," *Journal of Computational and Applied Mathematics*, vol. 166, no. 1, pp. 209–219, 2004.

[120] B.F. Spencer and L.A. Bergman, "On the numerical solution of the Fokker-Planck equation for nonlinear stochastic systems," *Nonlinear Dynamics*, vol. 4, no. 4, pp. 357–372, 1993.

[121] J.L. Brown, "Mean square truncation error in series expansions of random functions," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 1, pp. 28–32, 1960.

[122] A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, vol. 87, New York: Addison-Wesley, 1994.

[123] P. Langevin, "On the theory of Brownian motion," *CR Acad. Sci.(Paris)*, vol. 146, pp. 530–533, 1908.

[124] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized Algorithms for Analysis and Control of Uncertain Systems*, London: Springer-Verlag, 2005.

[125] A. Justel, D. Peña, and R. Zamar, "A multivariate Kolmogorov-Smirnov test of goodness of fit," *Statistics & Probability Letters*, vol. 35, no. 3, pp. 251–259, 1997.

[126] A. Dvoretzky, J. Kiefer, and J. Wolfowitz, "Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator," *The Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 642–669, 1956.

[127] R.H.C. Lopes, I. Reid, and P.R. Hobson, "The two-dimensional Kolmogorov-Smirnov test," in *The XI International Workshop on Advanced Computing and Analysis Techniques in Physics Research*, Nikhef, Amsterdam, Netherlands, April 2007.

[128] G. Steinbrecher and W.T. Shaw, "Quantile mechanics," *European Journal of Applied Mathematics*, vol. 19, no. 2, pp. 87–112, 2008.

[129] A. Halder and R. Bhattacharya, "Dispersion analysis in hypersonic flight during planetary entry using stochastic Liouville equation," *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 2, pp. 459–474, 2011.

[130] P. Dutta and R. Bhattacharya, "Hypersonic state estimation using Frobenius-Perron operator," *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 2, pp. 325–344, 2011.

[131] K. Pearson, "Contributions to the mathematical theory of evolution. II: Skew variation in homogeneous material," *Philosophical Transactions of the Royal Society of London. A*, vol. 186, pp. 343–414, 1895.

[132] M. Athans, "The role and use of the stochastic linear-quadratic-gaussian problem in control system design," *IEEE Transactions on Automatic Control*, vol.

16, no. 6, pp. 529–552, 1971.

[133] J.G. Michopoulos, C. Farhat, and J. Fish, "Modeling and simulation of multiphysics systems," *Journal of Computing and Information Science in Engineering (Transactions of the ASME)*, vol. 5, no. 3, pp. 198–213, 2005.

[134] L. Biegler, G. Biros, and O. Ghattas, *Large-Scale Inverse Problems and Quantification of Uncertainty*, vol. 707, West Sussex, UK: John Wiley & Sons, 2011.

[135] D.R. Herschbach, J. Avery, and O. Goscinski, *Dimensional Scaling in Chemical Physics*, Norwell, MA: Kluwer Academic Publishers, 1993.

[136] G. Chen, Z. Ding, S.B. Hsu, M. Kim, and J. Zhou, "Mathematical analysis of a Bohr atom model," *Journal of Mathematical Physics*, vol. 47, no. 2, pp. 022107_1–022107_23, 2006.

[137] F. Daum, "Dimensional interpolation for nonlinear filters," in *Proc. of SPIE*, vol. 5913, pp. 59131E_1–59131E_8.

[138] L. Rüschendorf, "The Wasserstein distance and approximation theorems," *Probability Theory and Related Fields*, vol. 70, no. 1, pp. 117–129, 1985.

[139] J.A. Cuesta and C. Matran, "Notes on the Wasserstein metric in Hilbert spaces," *The Annals of Probability*, vol. 17, no. 3, pp. 1264–1276, 1989.

APPENDIX A

ANALYSIS OF THE EIGENVALUES OF A POLYNOMIAL CHAOS APPROXIMATED SYSTEM

In this appendix, we present an eigenvalue analysis of a linear system with parametric uncertainty when polynomial chaos (PC) is used to approximate the random parameter. The main goal here is to show if the eigenvalues of a system are properly captured through using PC expansion.

We consider a linear system whose dynamics is given by the following equations

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \Delta & -3 \\ 4 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{A.1}$$

where $x = [x_1, x_2]^\top$ are the states. Δ is a random parameter having standard normal distribution. Our aim here is to get an approximation of the system in Eqn. (A.1) and to check if the deterministic system have the same eigenvalues as the actual system. We use Monte Carlo simulation for comparison of eigenvalues.

We start by expanding the states using finite term PC expansion, which are given by,

$$x_1(t, \Delta) = \sum_{i=1}^{N} x_{1i}(t)\phi_i(\Delta) \tag{A.2a}$$

$$x_2(t, \Delta) = \sum_{i=1}^{N} x_{2i}(t)\phi_i(\Delta) \tag{A.2b}$$

where $x_{1i}, x_{2i}, i = 1 \ldots, N$ are PC coefficients. $\phi_i(\Delta)$ are orthogonal polynomials which in this particular case are Hermite polynomials. The first ten Hermite polynomials are shown in Table IX.

Table IX.: The first ten Hermite polynomials

Number	**Polynomial**
ϕ_0	1
ϕ_1	Δ
ϕ_2	$\Delta^2 - 1$
ϕ_3	$\Delta^3 - 3\Delta$
ϕ_4	$\Delta^4 - 6\Delta^2 + 3$
ϕ_5	$\Delta^5 - 10\Delta^3 + 15\Delta$
ϕ_6	$\Delta^6 - 15\Delta^4 + 45\Delta^2 - 15$
ϕ_7	$\Delta^7 - 21\Delta^5 + 105\Delta^3 - 105\Delta$
ϕ_8	$\Delta^8 - 28\Delta^6 + 210\Delta^4 - 420\Delta^2 + 105$
ϕ_9	$\Delta^9 - 36\Delta^7 + 378\Delta^5 - 1260\Delta^3 + 945\Delta$
ϕ_{10}	$\Delta^{10} - 45\Delta^8 + 630\Delta^6 - 3150\Delta^4 + 4725\Delta^2 - 945$

Let

$$A = \begin{pmatrix} \int_{\mathbf{D}_\Delta} \Delta\phi_1^2(\Delta)p(\Delta)d\Delta & \cdots & \int_{\mathbf{D}_\Delta} \Delta\phi_1(\Delta)\phi_N(\Delta)p(\Delta)d\Delta \\ \vdots & \ddots & \vdots \\ \int_{\mathbf{D}_\Delta} \Delta\phi_N(\Delta)\phi_1(\Delta)p(\Delta)d\Delta & \cdots & \int_{\mathbf{D}_\Delta} \Delta\phi_N^2(\Delta)p(\Delta)d\Delta \end{pmatrix},$$

and,

$$G = \begin{pmatrix} \int_{\mathbf{D}_\Delta} \phi_1^2(\Delta)p(\Delta)d\Delta & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \int_{\mathbf{D}_\Delta} \phi_N^2(\Delta)p(\Delta)d\Delta \end{pmatrix},$$

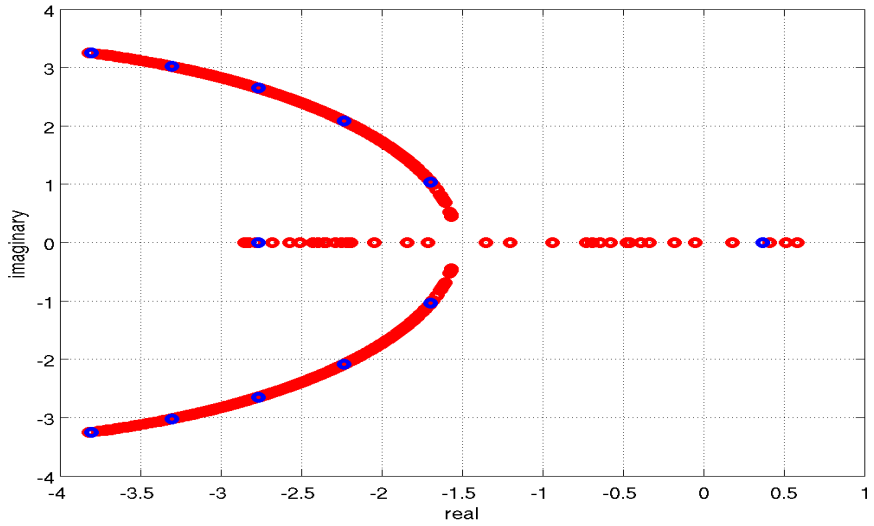
where $\mathbf{D}_\Delta$ is the domain of Δ and $p(\Delta)$ is the standard normal density. After applying PC to the dynamical system and taking inner products with respect to the basis functions, the PC dynamical system for Eqn. (A.1) is given by,

$$\frac{d}{dt}\begin{pmatrix} x_{11} \\ \vdots \\ x_{1N} \\ x_{21} \\ \vdots \\ x_{2N} \end{pmatrix} = \begin{pmatrix} G^{-1}A & \vdots & -3I \\ \cdots & & \cdots \\ 4I & \vdots & -5I \end{pmatrix} \begin{pmatrix} x_{11} \\ \vdots \\ x_{1N} \\ x_{21} \\ \vdots \\ x_{2N} \end{pmatrix}. \tag{A.3}$$

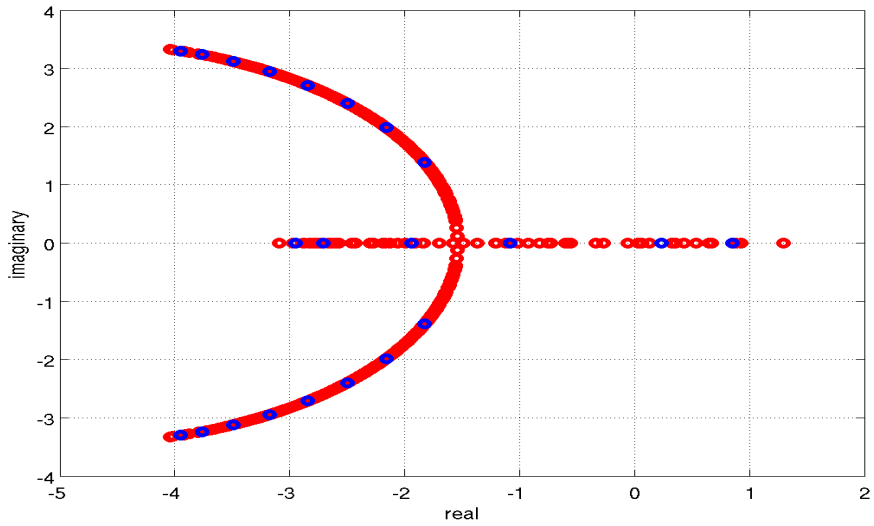
We plot the eigenvalues for the dynamical system in Eqn. (A.1) by varying Δ between its 3σ limits and overlay the eigenvalues of the PC dynamical system in Eqn. (A.3) on it. Figure 60 shows distribution of eigenvalues for the system in a complex plane, when 5 terms in the PC expansion are taken. It can be seen that the eigenvalues of PC dynamics match the eigenvalues of the actual dynamical system. Next, we increase the number of terms to $N = 10$. We see that the PC dynamical system is able to capture the eigenvalue distribution better than the previous case. We can conclude that, number of eigenvalues represented by PC in the complex plane

increases as N is increased. Hence, as $N \to \infty$ the PC dynamical system would completely represent the eigenvalue distribution of the actual dynamics.

In Fig. 61 we plot the PDF of the eigenvalues in the complex plane with the color denoting the corresponding density. It can be seen that the densities are symmetric about the real line, which is because complex conjugate eigenvalues have same density. We can not infer the density information from PC approximation. One of the drawbacks of using the PC to get eigenvalues of a linear system is that we do not have any PDF information after applying PC.



(a) $N = 5$.



(b) $N = 10$.

Fig. 60.: Locations of eigenvalues of the system in Eqn. (A.1), when Monte Carlo approximation is used (red circles) and when PC approximated dynamics in Eqn. (A.3) is used (blue).

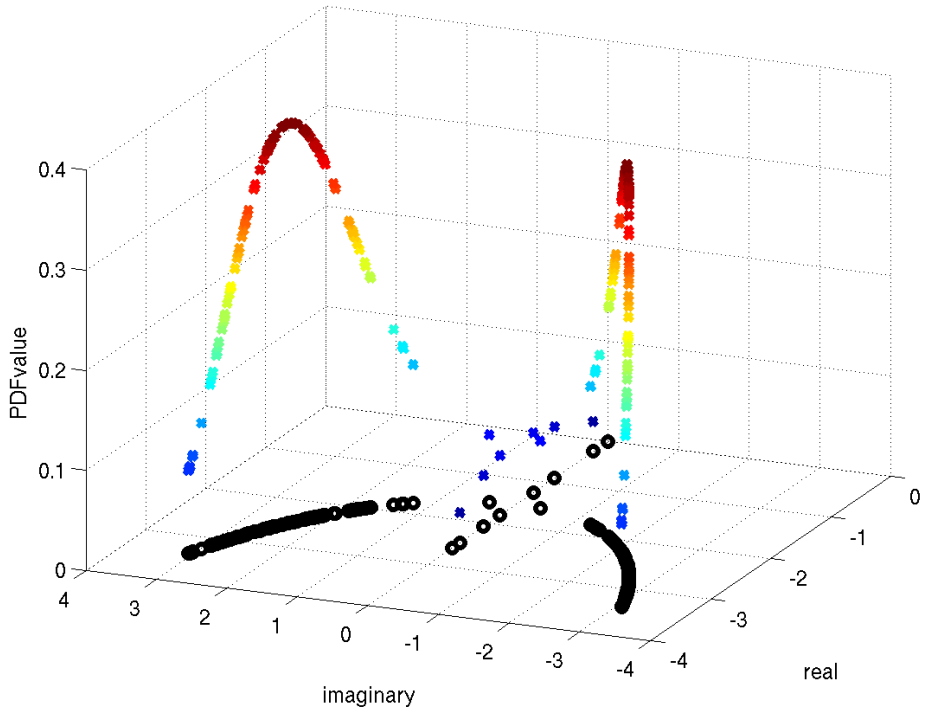


Fig. 61.: PDF of the eigenvalue distribution on the complex plane. Blue represents low probability regions and red is high probability region. Eigenvalues represented in black.

APPENDIX B

COMPARISON OF FROBENIUS-PERRON BASED ESTIMATION TO PARTICLE FILTERS

In this appendix, we compare the FP operator based estimation method to particle filters, by investigating the posterior PDF obtained by the two methods. We use a linear system with Gaussian uncertainty for comparison, as we have the exact representation of evolving sequence of posterior PDFs through Kalman filter. We compute the Wasserstein metric [138] of the two posteriors from the posterior obtained from Kalman filtering.

The Wasserstein metric is a metric comparing the distance between the shapes of two distributions. Let (M, d) be a metric space. For $p \geq 1$, let $P_p(M)$ denote the collection of all probability measures μ on M with finite p^{th} moment, i.e. for some x_0 in M,

$$\int_M d(x, x_0)^p d\mu(x) < \infty$$

Then the p^{th} Wasserstein distance between two probability measures μ and ν in $P_p(M)$ is defined as

$$\mathcal{W}_p(\mu, \nu) = \left(\inf_{\gamma \in \Gamma(\mu,\nu)} \int_{M \times M} d(x, y)^p \, \mathrm{d}\gamma(x, y) \right)^{1/p}, \tag{B.1}$$

where, $\Gamma(\mu, \nu)$ is the set of all pairing of measures (μ, ν) on $M \times M$. Important property of this metric is that it is commutative i.e. $\mathcal{W}_p(\mu, \nu) = \mathcal{W}_p(\nu, \mu)$. For the sake of brevity detailed discussion about the Wasserstein metric has been omitted here and can be found in ref. [139]. If both μ and ν are Gaussian measures, and if we consider $p = 2$, requiring the 2^{nd} moment to be finite, then there is an analytical

expression of Wasserstein metric given by,

$$W_2(\mu,\nu) = \sqrt{||\mu_1 - \mu_2||^2 + \mathrm{tr}(\Sigma_1) + \mathrm{tr}(\Sigma_2) - 2\,\mathrm{tr}(\sqrt{\Sigma_1}\Sigma_2\sqrt{\Sigma_1})^{\frac{1}{2}}}, \tag{B.2}$$

where μ_1, μ_2, Σ_1 and Σ_2 are the means and covariances of each Gaussian measure, respectively. The Wasserstein metric can be normalized by dividing $W_2(\mu,\nu)$ in Eqn. (B.2) by the diameter of the Ω space.

Let us consider the linear system given by the following dynamical equations.

$$\dot{x}(t) = \begin{pmatrix} -0.05 & 0 \\ 0 & -0.05 \end{pmatrix} x(t), \tag{B.3}$$

where $x(t) \in \mathbb{R}^2$ are the states. We consider a scalar measurement model given by,

$$\tilde{y}_k = [1 \quad 1]\, x_k + v_k, \tag{B.4}$$

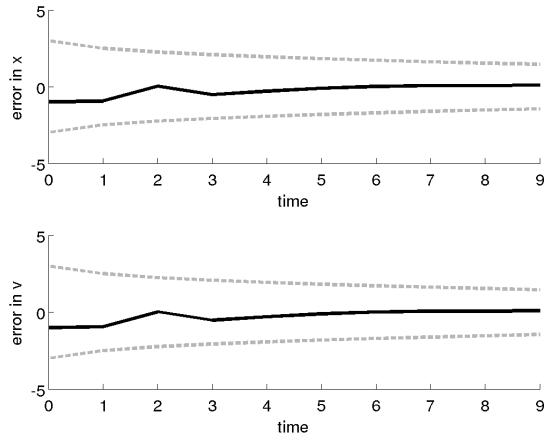
where $\tilde{y}_k$ are measurements coming at discrete times, and v_k is a zero mean delta correlated discrete Gaussian noise with autocorrelation R.

We assume that the system has Gaussian initial condition uncertainty with mean and covariance as,

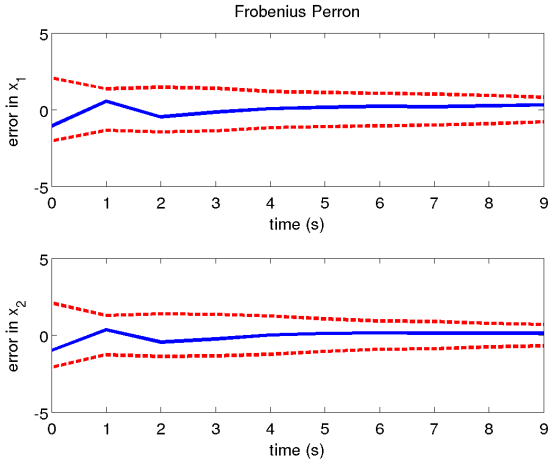
$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

respectively. The true system is assumed to have initial states $[2, 2]$, hence we start with an initial error in estimation. We fix the measurement update interval with each measurement coming every 1s. We run the estimation algorithms for 10s.

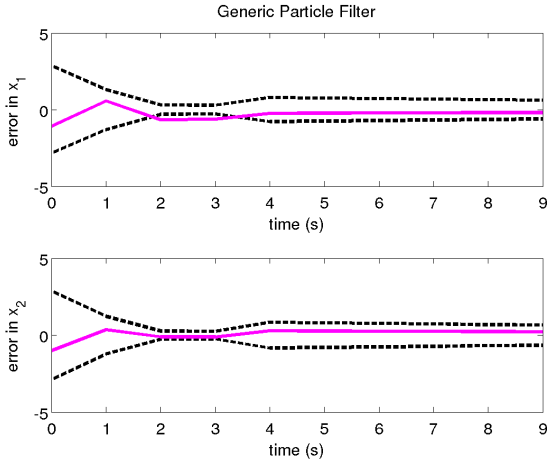
Initially, we fix the value of R to 2. Figure 62 shows plots for estimation error for the Kalman filter, FP operator based filter and particle filter. We can see that performance of FP operator based estimator is similar to that of the Kalman filter, in this case.



(a) Kalman filter.



(b) FP operator based estimator.



(c) Particle filter.

Fig. 62.: Plots for estimation error and $\pm 3\sigma$ limits for a) Kalman filter b) FP operator based estimator c) particle filter.

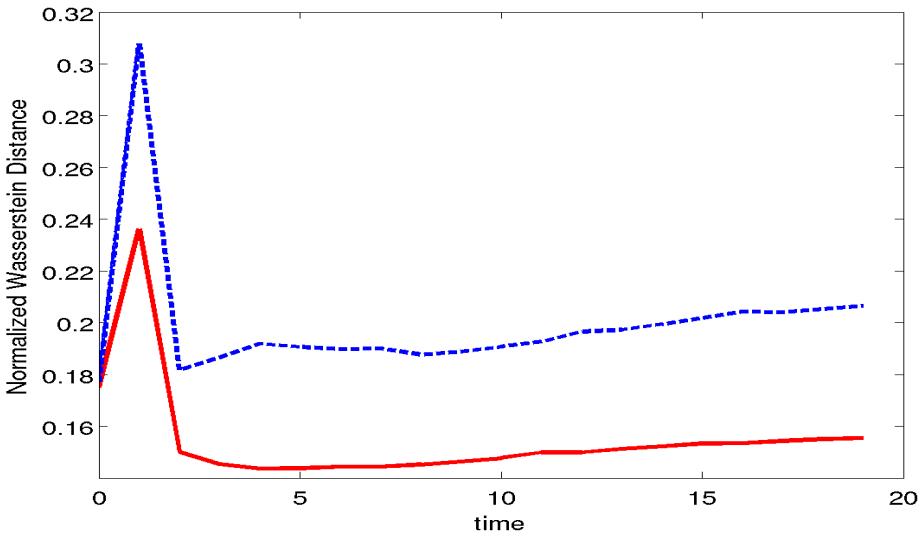


Fig. 63.: Normalized Wasserstein distance between posterior PDFs of Kalman filter and FP operator based estimator (solid line) and Kalman filter and particle filter (dashed line). The value of $R = 1/2$.

Figure 63 shows the plot for normalized Wasserstein metric for posterior PDFs, generated by particle filter and FP operator based estimator, from the posterior PDF of the Kalman filter when $R = 1/2$. The propagation time is fixed to 20s. We can see that the posterior PDF of the FP operator based estimator is closer to Kalman filter's posterior than the particle filter. If autocorrelation of noise is increased to $R = 2$, the difference in performance of the two estimators become even clearer, as it can be seen in Fig. 64.

However, if the value of R is increased the estimators perform equally. Figure 65 shows the Wasserstein distance between posterior PDFs of Kalman filter and FP operator based estimator and Kalman filter and particle filter, when value of $R = 4$. It can be seen that the distance from Kalman filter's posterior PDF increases for both filters, even if FP operator is nearer than particle filter.

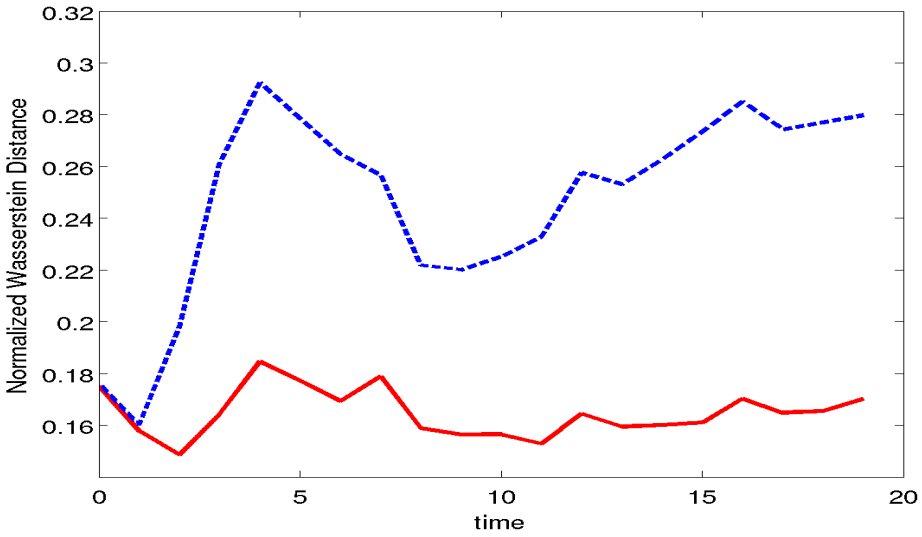


Fig. 64.: Normalized Wasserstein distance between posterior PDFs of Kalman filter and FP operator based estimator (solid line) and Kalman filter and particle filter (dashed line). The value of $R = 2$.

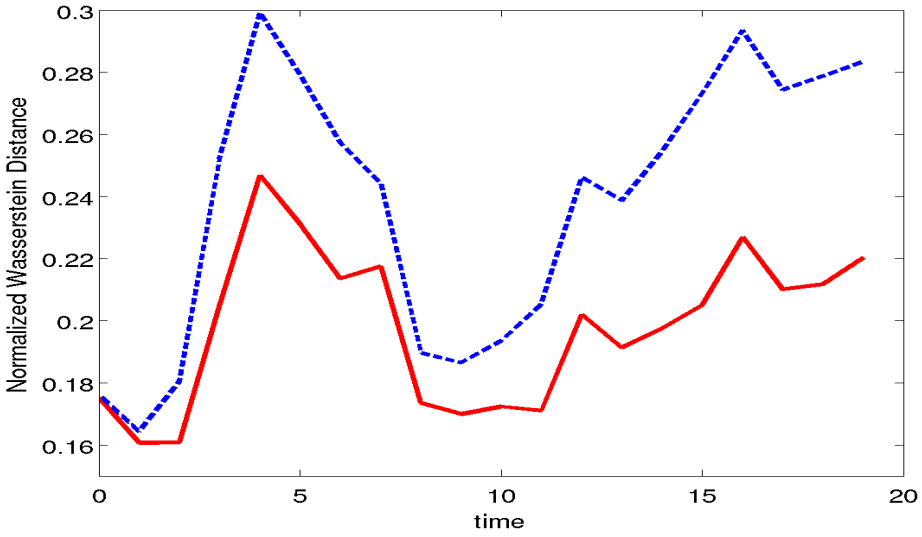


Fig. 65.: Normalized Wasserstein distance between posterior PDFs of Kalman filter and FP operator based estimator (solid line) and Kalman filter and particle filter (dashed line). The value of $R = 4$.

APPENDIX C

TENSOR CALCULATIONS

This appendix shows the tensor calculations for for finding the prior moments of the system in Eqn. (3.10) & Eqn. (3.11). After applying gPC to the system, the gPC dynamical system is given by Eqn. (3.14a) & Eqn. (3.14b). To get the prior higher order moments, we need to evaluate tensors which are given in section 3 in chapter III. The tensor calculations for the given problem are as follows.

$$P^{xv} = \begin{pmatrix} M^{3-}_{111} + M^{3-}_{122} \\ M^{3-}_{211} + M^{3-}_{222} \end{pmatrix} - \begin{pmatrix} M^{1-}_{1} M^{2-}_{11} + M^{1-}_{1} M^{2-}_{22} \\ M^{1-}_{2} M^{2-}_{11} + M^{1-}_{2} M^{2-}_{22} \end{pmatrix}, \tag{C.1a}$$

$$P^{vv} = R + M^{4-}_{1111} + M^{4-}_{2222} + 2M^{4-}_{1122} - 2\left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right)\left(M^{2-}_{11} + M^{2-}_{22}\right) + \left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right)^2, \tag{C.1b}$$

$$P^{xvv} = \begin{pmatrix} M^{5-}_{11111} + 2M^{5-}_{11122} + M^{5-}_{12222} \\ M^{5-}_{21111} + 2M^{5-}_{11222} + M^{5-}_{22222} \end{pmatrix} - 2\begin{pmatrix} M^{3-}_{111} + M^{3-}_{122} \\ M^{3-}_{211} + M^{3-}_{222} \end{pmatrix}\left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right) - \begin{pmatrix} M^{1-}_{1} \\ M^{1-}_{2} \end{pmatrix}(M^{4-}_{1111} + 2M^{4-}_{1122} + M^{4-}_{2222}) + 2\begin{pmatrix} M^{1-}_{1} \\ M^{1-}_{2} \end{pmatrix}\left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right)\left(M^{2-}_{11} + M^{2-}_{22}\right) - R\begin{pmatrix} M^{1-}_{1} \\ M^{1-}_{2} \end{pmatrix}, \tag{C.1c}$$

$$P^{xxv} = \begin{pmatrix} M^{4-}_{1111} + M^{4-}_{1122} & M^{4-}_{1112} + M^{4-}_{1222} \\ M^{4-}_{1112} + M^{4-}_{1222} & M^{4-}_{2211} + M^{4-}_{2222} \end{pmatrix} - \begin{pmatrix} M^{3-}_{111} + M^{3-}_{122} \\ M^{3-}_{211} + M^{3-}_{222} \end{pmatrix}\begin{pmatrix} M^{1-}_{1} \\ M^{1-}_{2} \end{pmatrix}^T - \begin{pmatrix} M^{2-}_{11} & M^{2-}_{12} \\ M^{2-}_{21} & M^{2-}_{22} \end{pmatrix}\left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right) - \begin{pmatrix} M^{1-}_{1} \\ M^{1-}_{2} \end{pmatrix}\begin{pmatrix} M^{3-}_{111} + M^{3-}_{122} \\ M^{3-}_{211} + M^{3-}_{222} \end{pmatrix}^T + \begin{pmatrix} M^{1-}_{1} \\ M^{1-}_{2} \end{pmatrix}\begin{pmatrix} M^{1-}_{1} \\ M^{1-}_{2} \end{pmatrix}^T\left(M^{2-}_{11} + M^{2-}_{22}\right) + \begin{pmatrix} M^{1-}_{1} \\ M^{1-}_{2} \end{pmatrix}\begin{pmatrix} M^{1-}_{1} \\ M^{1-}_{2} \end{pmatrix}^T\left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right)^2, \tag{C.1d}$$

$$P^{vvv} = M^{6-}_{111111} + 3M^{6-}_{111122} + 3M^{6-}_{112222} + M^{6-}_{222222} - 3\left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right)(M^{4-}_{1111} + 2M^{4-}_{1122} + M^{4-}_{2222}) - 3\left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right)R - 3\left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right)^2(M^{2-}_{11} + M^{2-}_{22}) - \left((M^{1-}_{1})^2 + (M^{1-}_{2})^2\right)^3. \tag{C.1e}$$

Here, M^{i-} are the prior moments which are calculated from the gPC coefficients. The superscript is the order of the moment and the subscript is the element of the corresponding moment tensor. For example M^{6-}_{112222} represents $(1,1,2,2,2,2)^{th}$ member of the 6^{th} order moment, which happens to be a 6^{th} order tensor.

VITA

Parikshit Dutta received his Bachelor of Technology degree in manufacturing science and engineering, and Master of Technology in industrial engineering and management, both in June 2006, as a part of dual degree program, from Indian Institute of Technology, Kharagpur, India, . He received his Ph.D. in aerospace engineering from Texas A&M University in August 2011. His research interests lie in the fields of uncertainty quantification of dynamical systems, estimation theory and sensitivity analysis, Monte Carlo methods for high dimensional systems, and Bayesian inferential techniques.

He can be reached via email at pdutta.kgp@gmail.com or at,
Department of Aerospace Engineering,
H.R. Bright Building, Rm. 701, Ross Street - TAMU 3141,
College Station TX 77843-3141.