Polynomial Chaos Based Method for State and Parameter Estimation

by

Reza Madankan

A thesis submitted to the Faculty of the Graduate School of the
State University of New York at Buffalo
in partial fulfillment of the requirements for the degree of
Master of Science
Department of Mechanical and Aerospace Engineering

Acknowledgement

It is a pleasure to thank those who made this thesis possible. My advisor, Dr. Puneet Singla, and my co-advisors Dr. Tarunraj Singh and Dr. Peter Scott for introducing me to the amazing field of estimation and their invaluable suggestions and guidance throughout my research work. Finally, I would like to thank my family for who always supported me.

This work was supported by the U.S. National Science Foundation (CMMI-0908403 and CMMI-1054759). Any findings, opinions and recommendations presented in this material are those of the author and do not necessarily reflect the view of NSF.

Reza Madankan
Fall 2011

Contents

Abstract

Two new recursive approaches have been developed to provide accurate estimates for posterior moments of both parameters and system states while making use of the generalized Polynomial Chaos (gPC) framework for uncertainty propagation. The main idea of the generalized polynomial chaos method is to expand random state and input parameter variables involved in a stochastic differential/difference equation in a polynomial expansion. These polynomials are associated with the prior pdf for the input parameters. Later, Galerkin projection is used to obtain deterministic system of equations for the expansion coefficients. The first proposed approach (gPC-Bayes) provides means to update prior expansion coefficients by constraining the polynomial chaos expansion to satisfy the desired number of posterior moment constraints derived from the Bayes' rule. The second proposed approach makes use of the minimum variance formulation to update polynomial chaos expansion coefficients. The main advantage of proposed methods is that they not only provide point estimate for the state and parameters but they also provide statistical confidence bounds associated with these estimates. Numerical experiments involving four benchmark problems are considered to illustrate the effectiveness of the proposed ideas.

Keywords:
Parameter Estimation, State Estimation, Inverse Problem, Generalized Polynomial Chaos, Bayes' Theorem, Method of Moments

Chapter 1

Introduction

Numerous fields of science and engineering require the study of the relevant stochastic dynamic system since mathematical models used to represent physical processes or engineering systems have errors and uncertainties associated with them. The error inherent in any mathematical model prediction can be due to the result of model truncation, errors in model parameters, errors in the inputs to the system and errors in initial conditions. These uncertainties cause overall accuracy of computations to degrade as the model states evolve. To alleviate this problem, assimilating the available observation data to correct and refine the model forecast in order to reduce the associated uncertainties is a logical improvement over purely model-based prediction. However, sensor model and data inaccuracies can lead to imprecise measurement data which could lead to inaccurate estimates. Hence, the optimal solution should be a weighted mixture of model forecast and observation data. This approach had its birth with the development of the Kalman Filter [1].

Kalman Filter (KF) is the optimal Bayesian estimator for linear systems with initial condition and measurement errors assumed to be Gaussian. However, the performance of the Kalman filter can deteriorate appreciably due to model parameter uncertainty [2, 3, 4]. The sensitivity of the KF to parametric modeling errors has led to the development of several robust filtering approaches; robust in the sense that they attempt to limit, in certain ways, the effect of parameter uncertainties on the overall filter performance. Various approaches to state-space estimation in this regard [5] have focused on $\mathcal{H}_\infty$ filtering [6, 7], set-valued estimation [8, 9], and guaranteed cost designs [8, 10]. Alternatively, when the model parameters are uncertain, the estimation is carried out through the simultaneous

estimation of states and parameters (also viewed as states), which results in a nonlinear filtering problem even for otherwise linear systems [11]. Methods like the extended Kalman Filter (EKF) [2] or Unscented Kalman Filter (UKF) [12] have been used to estimate model parameters along with state estimates. In the EKF approach, the original nonlinear model is converted to a linearized model by using the jacobian of the nonlinear model about current state and parameter estimates. A major drawback of the EKF approach is that it results in poor performance when the state transition or observation models are highly nonlinear or even if state estimates are highly sensitive to parametric errors in case of a linear system. Unscented Kalman Filter (UKF) is one of the approaches which can be used to overcome this deficiency. UKF performs the estimation process by making use of a deterministic sampling technique known as the unscented transformation. Unscented transformation provides a set of sample points around the mean (known as σ-points) which are propagated through the nonlinear functions, from which the mean and covariance of the estimate are then recovered. This process results in a filter which captures the true mean and covariance better than the EKF.

Although both the EKF and UKF based filters are very popular for simultaneous state and parameter estimation problems, both methods are based upon very restrictive Gaussian error assumption for both parameter and state uncertainty. Clearly, the Gaussian assumption can work well for moderately nonlinear systems but it might not be appropriate at all for certain problems based upon the physical model. For example, Gaussian distribution is not an ideal distribution to represent errors in the spring coefficient which is a positive quantity. This necessitates the need for filters which can incorporate the knowledge about non-Gaussian uncertainty. Various researches have endeavored to exploit knowledge of statistics, dynamic systems and numerical analysis to develop nonlinear filtering techniques [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23] which cater to the various classes of state and parameter estimation problems. For low-order nonlinear systems, the Particle Filter (PF) [22, 23] has been gaining increasing attention. However, Daum in his seminal work[24] discusses that various factors like *volume of state space in which conditional pdf is non-vanishing, rate of decay of the conditional pdf in state space, stationarity of the problem, analytical structure of the problem (e.g. linear dynamics, bilinear dynamics, unimodal pdf, etc.), effective dimensionality of the problem, etc.* strongly affect the computational complexity and performance of the particle filter and argue that an efficient approach for general

nonlinear filtering should be based upon continuous probability density function (pdf) assumption [24].

For linear systems with parametric uncertainties, the multiple-model estimation [11] methods have been very popular. This method assumes the uncertain parameters belong to a discrete set. The uncertain parameter vector is quantized to a finite number of grid points with known prior probabilities. The state conditional mean and covariance are propagated for each model corresponding to a grid point using KF equations and the first two moments of system states are computed by a weighted average of the moments corresponding to various prior models. Furthermore, the prior probability values for parameter samples are also updated by making use of the Bayes' theorem. Although this method works well for linear systems and provides a mean estimate for both state and parameter, the performance of this method is strongly affected by number of parameter samples like any sampling algorithm such as PF [25]. A detailed review on classical approaches applied in online parameter estimation can be found in [26].

1.1 Polynomial Chaos Based Estimation

All the methods mentioned in previous section have some restrictions for application. As mentioned before, all the Kalman based filters, like KF, EKF, and UKF have a restrictive assumption about the distribution of the parameters and states. Also, application of PF encounters expensive computational cost for large number of samples applied during the estimation process. One of the proposed approaches to overcome these restrictions is generalized Polynomial Chaos (gPC) based estimation methods.

gPC is an extension of the polynomial chaos (PC) idea of Wiener [27]. The main principle of the polynomial chaos approach is to expand random variables using polynomial basis functions that are orthogonal with respect to the pdf of the parameters (Hermite polynomials for normally distributed parameters, Legendre for uniformly distribution, etc.), and transform stochastic equations into deterministic equations in higher dimensional projection space using Galerkin collocation. Xiu et al. [28] generalized the result of Cameron-Martin to various continuous and discrete distributions using orthogonal polynomials from the so called Askey-scheme [29]. This is popularly known as the gPC framework. The gPC based methods have emerged as powerful tools to propagate time-invariant parametric uncertainty

through an otherwise deterministic system of equations, to predict a distribution of outputs [27, 28, 30]. The gPC method can efficiently characterize the state uncertainty due to time-invariant random parameters having arbitrary probability distributions.

gPC has been used in different ways for parameter estimation problem also popular in literature as *inverse problem*. Blanchard *et al.* [31] combined gPC method with Extended Kalman Filter (EKF). In the framework of this approach, after application of gPC in finding the solution of forward problem, an (suboptimal) EKF is used to recalculate the polynomial chaos expansions for the uncertain states and the uncertain parameters using the Galerkin projection. In another similar work, a recursive approach which constructs a set of efficient algorithms based on combination of the gPC expansion and the Ensemble Kalman Filter (EnKF) has been proposed [32]. The key steps in the proposed approach involve solving the system of stochastic state equations via the gPC-based numerical methods (stochastic Galerkin or stochastic collocation) to gain efficiency, then sampling the gPC approximation of the stochastic solution with an arbitrarily large number of samples in order to reduce sampling errors. The drawback of this work is that like all Kalman based estimation approaches, they assume distribution of parameter of interest to be Gaussian.

The gPC approach in conjunction with maximum likelihood framework has been used successfully to find the estimates for parameters, but not states. Pence *et al.* [33] proposed a recursive algorithm based on gPC approach and maximum likelihood estimation to find the value of static parameters of linear or nonlinear stochastic dynamic systems, given the system's inputs. In this method, the gPC approach is used to propagate the uncertainty of the system through the forward dynamic model and a maximum likelihood function is composed based on measurement data and the gPC propagation of the system. Finally, the point estimate of parameter of interest at every time step is obtained by finding the corresponding realization of random variable which maximizes the likelihood function. The paper suggests two different approaches to solve the maximum likelihood problem at each time step to find the update of random variable, which are Gradient based optimization or random search approach. The main drawback of this approach is that it provides a point estimate rather than the posterior density function for parameters.

Recently, the gPC approach has also been used in a Bayesian framework to provide estimates for parameters in the context of solving inverse problems. Marzouk *et al.* [34] demonstrated an approach to evaluate posterior distribution of a parameter of interest by

combination of the gPC approach, Bayesian framework and Markov Chain Monte Carlo (MCMC) method. In this approach, the gPC-based stochastic Galerkin method is used to propagate prior uncertainty through the forward model. Then, application of gPC approximation of forward solution in evaluation of prior and the likelihood function in Bayesian framework leads to a posterior distribution of parameters of interest. Finally, MCMC sampling technique has been used to explore the obtained posterior distribution. In another research, Marzouk *et al.* [35] used a gPC stochastic collocation approach to construct posterior surrogates for efficient Bayesian inference in inverse problems, instead of the Galerkin method. This work also contains a rigorous error analysis of the gPC Bayesian inverse scheme. Convergence of the approximate posterior distribution to the true posterior distribution is established and its asymptotic convergence rate is obtained. Please note that both of these researches have been implemented as batch estimation approach.

gPC approach has also been used in Bayesian framework to find a point estimate of parameters of interest. Blanchard *et al.* [36] proposed an offline computational approach for parameter identification based on the application of the generalized polynomial chaos theory which leads to a point estimation. In the approach presented in this paper, gPC expansion theory is used to propagate the solution of the system between the measurement updates. Then, the point estimate of parameter of interest is obtained by maximizing the posterior probability density function which is expressed in terms of prior probability and likelihood function, by using Bayes' theorem.

Recently, the gPC expansion has also been used in Maximum Entropy framework for recursive estimation purposes. Dutta *et al.* [37], developed a nonlinear estimation algorithm based on the combination of gPC expansion theory, Maximum Entropy principle, and higher order moments updates. In this research, polynomial chaos theory is used to predict the evolution of uncertainty in the nonlinear random process. Then, higher order moment updates are used to estimate the posterior moments of the random process using a linear gain. Finally, posterior probability density function is approximated by a mixture of Gaussian kernels by using the maximum entropy principle, subjected to constraints defined by the posterior moments. Also, the basis functions applied in gPC expansions are reconstructed according to the obtained pdf, after each update. However similar to [4], they considered just state estimation in the presence of parametric or initial condition uncertainty. Furthermore, the approximation by Gaussian kernels require special tuning which can be cumbersome for

many real problems.

1.2 Present Work

In summary, the gPC expansion method has been successfully used to find point estimates by making use of maximum likelihood or maximum posteriori framework. However, most of these methods just provides a point estimate rather than a complete description of the posterior pdf for both state and parameter. Furthermore, it should be noted that all these methods are either applied to state or parameter estimation problem and also most of them are being applied as an offline estimation approach.

This thesis presents two new recursive approaches to provide estimates for posterior moments of both parameters and system states by making use of the gPC expansion and the Bayesian framework. The main advantage of proposed methods is that they not only provide point estimate for the state and parameters but they also provide statistical confidence bounds associated with these estimates described in terms of the posterior moments. Furthermore, these moments have been applied in the construction of posterior coefficients of the gPC expansion for both states and parameters.

The first objective of this approach is to quantify the effect of parametric and initial conditions uncertainty on the output of the mathematical model. The second objective of this work is to provide a *real-time* estimate for system states and parameters, together with quantitative measures of confidence in that estimate, while taking into account model and sensor inaccuracies. In addition, it must account for non-Gaussian parametric and initial condition uncertainty and must be robust in presence of low measurement data frequency. Finally, regarding its application in *real time* estimation, it should be computationally affordable. In the following chapters, after theoretical development of our approach, we analyze performance of this method on different problems, especially regarding these requirements.

The remainder of this document is structured as follows: In chapter 2 we briefly review the generalized polynomial chaos theory and its application to model stochastic differential equations. The efficiency of gPC theory in quantifying the effect of parametric and initial conditions uncertainty has been shown by some numerical examples. In chapter 3, we describe the problem statement and formulation of estimation process by using Bayes rule and minimum variance estimator. Also, detailed formulations of measurement update process

are developed. Next, we illustrate the efficacy of this approach by some numerical examples in chapter 4. Finally, conclusion and discussion of the results are mentioned in chapter 5.

Chapter 2

Generalized Polynomial Chaos

This chapter presents the mathematical details for the polynomial chaos methodology to examine the effects of input parameter uncertainty on the forward model outcome.

A mechanism to represent the uncertainty is necessary before the model data and the sensed data can be integrated in an efficient and consistent manner. Probabilistic means of representing uncertainties has been explored extensively and provides the greatest wealth of knowledge which will be exploited in this work. In the standard dynamic model, the state variables are assumed to be a deterministic quantity. Instead of solving for the point estimates for state variables, we are interested in probability distribution for their values due to uncertainty in input parameters, initial conditions and random input. Thus, the system states are assumed to be a *random vector*, $\mathbf{x}(t)$, whose time evolution is given by the following stochastic differential equation:

$$\dot{\mathbf{x}}(t, \boldsymbol{\Theta}) = \mathbf{f}(t, \boldsymbol{\Theta}, \mathbf{x}, \mathbf{u}), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{2.1}$$

In this equation, $\boldsymbol{\Theta}$ represents uncertain but time-invariant system parameters and $\mathbf{u}$ represents deterministic forcing terms. The nominal initial state estimates are given by $\mathbf{x}_0$, which may also be uncertain. The total uncertainty associated with the state vector $\mathbf{x}(t_k)$ is characterized by the probability distribution function (pdf) $p(t_k, \mathbf{x}_k, \boldsymbol{\Theta})$. A key idea of this work is to *replace the time evolution of state vector* $\mathbf{x}_k$ *by the time evolution of the pdf* $p(t_k, \mathbf{x}_k, \boldsymbol{\Theta})$ as illustrated in Fig. 2.1. By computing full probability density functions, we can better monitor the space-time evolution of uncertainty, represent multi-modal distributions, incorporate complex prior models, and exploit Bayesian belief propagation.

Several approximate techniques exist in the literature to approximate the state pdf evolu-

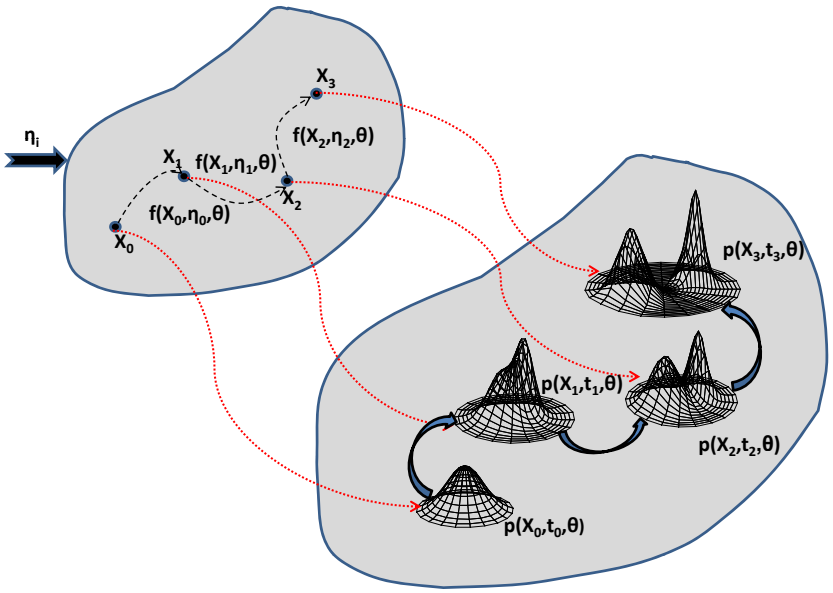


Figure 2.1: State and pdf transition

tion [38, 39], the most popular being Monte Carlo (MC) methods [40], Gaussian closure [41], Equivalent Linearization [42], and Stochastic Averaging [43, 44]. In addition, a Gaussian Process approach to solve nonlinear stochastic differential equations has been proposed in Ref. [45]. Here the Kullback-Leibler divergence [46] between the true posterior and the Gaussian approximation is minimized after approximating the first two moments of the posterior, by local linearization. All of these algorithms except MC methods are similar in several respects, and are suitable only for linear or moderately nonlinear systems, because the effect of higher order terms can lead to significant errors. Monte Carlo methods require extensive computational resources and effort, and become increasingly infeasible for high-dimensional dynamic systems [24].

The next section discusses the generalized Polynomial Chaos (gPC) method for solving the time evolution of state pdf for systems that include initial condition and parametric uncertainty.

2.1 Generalized Polynomial Chaos, Theory and Methodology

The propagation of uncertainty due to time-invariant but uncertain input parameters can be approximated by a generalization of polynomial chaos (gPC). gPC is an extension of the homogenous chaos idea of Wiener [47] and involves a separation of random variables from deterministic ones in the solution algorithm for a stochastic differential equation. The random variables are expanded in a polynomial expansion. These polynomials are associated with the assumed pdf for the input variables (Hermite polynomials for normally distributed parameters, Legendre for uniformly distribution, etc [48]). Galerkin projection is used to generate a system of deterministic differential equations for the expansion coefficients.

2.1.1 Linear Systems

To describe the gPC process in detail, let us first consider a generic first order stochastic linear system:

$$\dot{\mathbf{x}}(t,\boldsymbol{\Theta}) = \mathbf{A}(\boldsymbol{\Theta})\mathbf{x}(t,\boldsymbol{\Theta}) + \mathbf{B}(\boldsymbol{\Theta})\mathbf{u}(t) \tag{2.2}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. $\mathbf{u} \in \mathbb{R}^{p\times 1}$ is vector of input signals and $\boldsymbol{\Theta} \in \mathbb{R}^{r}$ is a vector of uncertain system parameters which is a function of the random variable $\boldsymbol{\xi}$ with known probability distribution function (pdf) $p(\boldsymbol{\xi})$. It is assumed that the uncertain state vector $\mathbf{x}(t,\boldsymbol{\Theta})$ and system parameters A_{ij} and B_{ij} can be written as a linear combination of basis functions, $\phi_k(\boldsymbol{\xi})$, which span the stochastic space of random variable $\boldsymbol{\xi}$:

$$x_i(t,\boldsymbol{\xi}) = \sum_{k=0}^{N} x_{i_k}(t)\phi_k(\boldsymbol{\xi}) = \mathbf{x}_i^T(t)\boldsymbol{\Phi}(\boldsymbol{\xi}) \tag{2.3}$$

$$A_{ij}(\boldsymbol{\xi}) = \sum_{k=0}^{N} a_{ij_k}\phi_k(\boldsymbol{\xi}) = \mathbf{a}_{ij}^T\boldsymbol{\Phi}(\boldsymbol{\xi}) \tag{2.4}$$

$$B_{ij}(\boldsymbol{\xi}) = \sum_{k=0}^{N} b_{ij_k}\phi_k(\boldsymbol{\xi}) = \mathbf{b}_{ij}^T\boldsymbol{\Phi}(\boldsymbol{\xi}) \tag{2.5}$$

where $\boldsymbol{\Phi}(.) \in \mathbb{R}^N$ is a vector of polynomial basis functions orthogonal to the pdf $p(\boldsymbol{\xi})$ which can be constructed using the *Gram-Schmidt Orthogonalization Process*. Table 2.1 represents different types of polynomial basis functions corresponding to different distributions of random variable $\boldsymbol{\xi}$ [48].

The coefficients a_{ij_k} and b_{ij_k} are obtained by making use of following *normal equations*:

$$a_{ij_k} = \frac{\langle \mathbf{A}_{ij}(\boldsymbol{\Theta}(\boldsymbol{\xi})), \phi_k(\boldsymbol{\xi}) \rangle}{\langle \phi_k(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi}) \rangle} \tag{2.6}$$

$$b_{ij_k} = \frac{\langle \mathbf{B}_{ij}(\boldsymbol{\Theta}(\boldsymbol{\xi})), \phi_k(\boldsymbol{\xi}) \rangle}{\langle \phi_k(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi}) \rangle} \tag{2.7}$$

where $\langle u(\boldsymbol{\xi}), v(\boldsymbol{\xi}) \rangle = \int\limits_{\mathbb{R}^r} u(\boldsymbol{\xi}) v(\boldsymbol{\xi}) p(\boldsymbol{\xi}) d\boldsymbol{\xi}$ represents the inner product induced by pdf $p(\boldsymbol{\xi})$.

Please note that the total number of terms in gPC expansion (N) is determined by the chosen highest order of basis polynomials $\phi_k(\boldsymbol{\xi})$, denoted by l, and the dimension of the vector of uncertain parameter $\boldsymbol{\Theta}$, which is represented by m:

$$N = \begin{pmatrix} l+m \\ m \end{pmatrix} = \frac{(l+m)!}{m!l!} \tag{2.8}$$

Now, substitution of Eq. (2.3), Eq. (2.4) and Eq. (2.5) in Eq. (2.2) leads to:

$$\begin{aligned} \mathbf{e}_i(\boldsymbol{\xi}) = \sum_{k=0}^{N} \dot{x}_{i_k}(t)\phi_k(\boldsymbol{\xi}) - \sum_{j=1}^{n}(\sum_{k=0}^{N} a_{ij_k}\phi_k(\boldsymbol{\xi}))(\sum_{k=0}^{N} x_{i_k}(t)\phi_k(\boldsymbol{\xi})) \\ - \sum_{j=1}^{p}(\sum_{k=0}^{N} b_{ij_k}\phi_k(\boldsymbol{\xi}))u_j, \quad i = 1, 2, \cdots, n \end{aligned} \tag{2.9}$$

Eq. (2.9) represents the error of approximate gPC solution of Eq. (2.2) which contains $n(N+1)$ time-varying unknown coefficients $x_{i_k}(t)$. These unknown coefficients can be obtained by using Galerkin process, i.e., projecting the error of Eq. (2.2) onto space of basis functions $\phi_k(\boldsymbol{\xi})$.

Table 2.1: Correspondence of polynomial basis functions with their underlying random variables $\boldsymbol{\xi}$

Random Variable $\boldsymbol{\xi}$	basis polynomials $\boldsymbol{\phi}(.)$	Support
Gaussian	Hermite	$(-\infty, +\infty)$
Gamma	Laguerre	$[0, +\infty]$
Beta	Jacobi	$[a, b]$
Uniform	Legendre	$[a, b]$

$$\langle e_i(\mathbf{C}, \boldsymbol{\xi}), \phi_k(\boldsymbol{\xi}) \rangle = 0, \quad i = 1, 2, \cdots, n, \quad k = 1, 2, \cdots, N \tag{2.10}$$

This leads to following set of $n(N+1)$ *deterministic differential equations*:

$$\dot{\mathbf{x}}_{pc}(t) = \mathcal{A}\mathbf{x}_{pc}(t) + \mathcal{B}\mathbf{u}(t) \tag{2.11}$$

where $\mathbf{x}_{pc}(t) = \{\mathbf{x}_1^T(T),\ \mathbf{x}_2^T(T),\ \cdots, \mathbf{x}_n^T(T)\}$ is a vector of $n(N+1)$ unknown coefficients, $\mathcal{A} \in \mathbb{R}^{n(N+1) \times n(N+1)}$ and $\mathcal{B} \in \mathbb{R}^{n(N+1) \times p}$.

Let P and T_k, for $k = 0, 1, 2, \cdots, N$, denote the inner product matrices of the orthogonal polynomials defined as follows:

$$P_{ij} = \langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle, \quad i, j = 0, 1, 2, \cdots, N \tag{2.12}$$

$$T_{k_{ij}} = \langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi}) \rangle, \quad i, j = 0, 1, 2, \cdots, N \tag{2.13}$$

Then $\mathcal{A}$ and $\mathcal{B}$ can be written as an $n(N+1) \times n(N+1)$ block-diagonal matrix, each on-diagonal block being an $(N+1) \times (N+1)$ matrix. The matrix $\mathcal{A}$ consists of blocks $\mathcal{A}_{ij} \in \mathbb{R}^{(N+1)\times(N+1)}$:

$$\mathcal{A}_{ij} = A_{ij}P, \quad i, j = 1, 2, \cdots, n \tag{2.14}$$

if matrix $\mathbf{A}$ is not uncertain, else, it is given by:

$$\mathcal{A}_{ij}(k, :) = \mathbf{a}_{ij}^T T_k, \quad i, j = 1, 2, \cdots, n \tag{2.15}$$

The matrix $\mathcal{B}$ consists of blocks $\mathcal{B}_{ij} \in \mathbb{R}^{(N+1)\times 1}$:

$$\mathcal{B}_{ij} = P\mathbf{b}_{ij} \quad i = 1, 2, \cdots, n, \quad j = 1, 2, \cdots, p \tag{2.16}$$

Eq. (2.3) along with Eq. (2.11) define the uncertain state vector $\mathbf{x}(t, \xi)$ as a function of random variable $\boldsymbol{\xi}$ and can be used to compute any order moment or cumulant of a function of uncertain state variable. For example, the first two moments for state vector $\mathbf{x}(t)$ can be written as:

$$\mathcal{E}[x_i(t)] = x_{i_1}(t), \quad i = 1, \cdots, n \tag{2.17}$$

$$\mathcal{E}[x_i(t)x_j(t)] = \sum_{k=0}^{N} x_{i_k}(t)x_{j_k}(t), \quad i, j = 1, \cdots, n \tag{2.18}$$

2.1.2 Nonlinear Systems

In this section, we extend the gPC process to propagate the state uncertainty for a generic nonlinear system given by

$$\dot{\mathbf{x}}(t,\boldsymbol{\Theta}) = \mathbf{f}(t,\boldsymbol{\Theta},\mathbf{x},\mathbf{u}), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{2.19}$$

where $\mathbf{u}(t)$ is the input of dynamic system at time t, $\mathbf{x}(t,\boldsymbol{\Theta}) = [x_1(t,\boldsymbol{\Theta}), x_2(t,\boldsymbol{\Theta}), \cdots, x_n(t,\boldsymbol{\Theta})]^T \in \mathbb{R}^n$ represents the stochastic system state vector, and uncertain parameter vector $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_m]^T \in \mathbb{R}^m$ is assumed to be time invariant and function of a random vector $\boldsymbol{\xi} = [\xi_1, \xi_2, \cdots, \xi_m]^T \in \mathbb{R}^m$ defined by a pdf $p(\boldsymbol{\xi})$ over the support Ω. Please note that $\mathbf{f}(t,\boldsymbol{\Theta},\mathbf{x},\mathbf{u})$ can be a nonlinear function in general.

Once again the gPC expansion for the state vector $\mathbf{x}$ and uncertain parameter $\boldsymbol{\Theta}$ can be written as:

$$x_i(t,\boldsymbol{\Theta}) = \sum_{k=0}^{N} x_{i_k}(t)\phi_k(\boldsymbol{\xi}) = \mathbf{x}_i^T(t)\boldsymbol{\Phi}(\boldsymbol{\xi}) \Rightarrow \mathbf{x}(t,\boldsymbol{\xi}) = \mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}) \tag{2.20}$$

$$\theta_i(\boldsymbol{\xi}) = \sum_{k=0}^{N} \theta_{i_k}\phi_k(\boldsymbol{\xi}) = \boldsymbol{\theta}_i^T\boldsymbol{\Phi}(\boldsymbol{\xi}) \Rightarrow \boldsymbol{\Theta}(t,\boldsymbol{\xi}) = \boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}) \tag{2.21}$$

where, $\mathbf{X}_{pc}$ and $\boldsymbol{\Theta}_{pc}$ are matrices composed of coefficients of gPC expansion for state $\mathbf{x}$ and parameter $\boldsymbol{\Theta}$, respectively. Similar to the linear case, coefficients θ_{i_k} are obtained by making use of following *normal equations*:

$$\theta_{i_k} = \frac{\langle \theta_i(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi}) \rangle}{\langle \phi_k(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi}) \rangle} \tag{2.22}$$

Now, substitution of Eq. (2.20) and Eq. (2.21) into Eq. (2.19), leads to:

$$\mathbf{e}_i(\mathbf{X}_{pc},\boldsymbol{\xi}) = \sum_{k=0}^{N} \dot{x}_{i_k}(t)\phi_k(\boldsymbol{\xi}) - \mathbf{f}_i(t,\mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}),\boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}),\mathbf{u}), \quad i = 1,2,\cdots,n \tag{2.23}$$

From Eq. (2.10), $n(N+1)$ time-varying coefficients x_{i_k} can be obtained using the Galerkin process, i.e. projecting the error captured in Eq. (2.23) onto space of basis functions $\phi_k(\boldsymbol{\xi})$.

For polynomial or rational state nonlinearity, the Galerkin process will lead to a set of $n(N+1)$ *nonlinear deterministic differential equations*. For non-polynomial nonlinearity such as transcendental or exponential functions, difficulties may arise during the computation of projection integrals of Eq. (2.10). To overcome this, in the nonlinear case polynomial chaos quadrature (PCQ) technique will be used.

2.2 Polynomial Chaos Quadrature

To manage the non-polynomial nonlinearity difficulties in polynomial chaos integration, Dalbey *et al.* have proposed a formulation [49] known as Polynomial Chaos Quadrature (PCQ). PCQ replaces the projection step of the gPC with numerical quadrature. The resulting method can be viewed as a MC-like evaluation of system equations, but with sample points selected by quadrature rules. To illustrate this, consider Eq. (2.19), which by substitution of Eq. (2.20) and Eq. (2.21) can be written as:

$$\sum_{k=0}^{N} \dot{x}_{i_k}(t)\phi_k(\boldsymbol{\xi}) - \mathbf{f}_i(t, \mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}), \boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}), \mathbf{u}) = 0, \quad i = 1, \cdots, n \tag{2.24}$$

The projection step of PC yields:

$$\sum_{k=0}^{N} \langle \phi_k(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle \dot{x}_{i_k} - \langle \mathbf{f}_i(t, \mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}), \boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}), \mathbf{u}), \phi_j(\boldsymbol{\xi}) \rangle = 0 \quad i = 1, \cdots, n, \quad j = 0, \cdots, N \tag{2.25}$$

In the case which $f(t, \mathbf{x}, \boldsymbol{\Theta}, \mathbf{u})$ is linear, it is possible to evaluate projection integrals of Eq. (2.25) analytically. More generally, the starting point of PCQ methodology is to replace the exact integration with respect to $\boldsymbol{\xi}$ by numerical integration. The familiar Gauss quadrature method is a suitable choice for most cases. This yields:

$$\langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle = \int \phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi})p(\boldsymbol{\xi})d\boldsymbol{\xi} \simeq \sum_{q=1}^{M} w_q \phi_i(\boldsymbol{\xi}_q)\phi_j(\boldsymbol{\xi}_q) \tag{2.26}$$

$$\langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi})\phi_k(\boldsymbol{\xi}) \rangle = \int \phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi})\phi_k(\boldsymbol{\xi})p(\boldsymbol{\xi})d\boldsymbol{\xi} \simeq \sum_{q=1}^{M} w_q \phi_i(\boldsymbol{\xi}_q)\phi_j(\boldsymbol{\xi}_q)\phi_k(\boldsymbol{\xi}_q) \tag{2.27}$$

$$\begin{aligned} \langle \mathbf{f}_i(t, \mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}), \boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}), \mathbf{u}), \phi_j(\boldsymbol{\xi}) \rangle &= \int \mathbf{f}_i(t, \mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}), \boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}), \mathbf{u})\phi_j(\boldsymbol{\xi})p(\boldsymbol{\xi})d\boldsymbol{\xi} \\ &\simeq \sum_{q=1}^{M} w_q \mathbf{f}_i(t, \mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}_q), \boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}_q), \mathbf{u})\phi_j(\boldsymbol{\xi}_q) \end{aligned} \tag{2.28}$$

where M is the number of quadrature points used. Substitution of aforementioned approximation of stochastic integral in Eq. (2.25) and interchanging summation and differentiation leads to

$$\frac{d}{dt} \sum_{q=1}^{M} \sum_{k=0}^{N} w_q \phi_j(\boldsymbol{\xi}_q)\phi_k(\boldsymbol{\xi}_q)x_{i_k} - \sum_{q=1}^{M} w_q \mathbf{f}_i(t, \mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}_q), \boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}_q), \mathbf{u})\phi_j(\boldsymbol{\xi}_q) = 0 \tag{2.29}$$

which can be simplified as:

$$\frac{d}{dt}\sum_{q=1}^{M}\phi_j(\boldsymbol{\xi}_q)x_i(t,\boldsymbol{\xi}_q)w_q - \sum_{q=1}^{M} w_q \mathbf{f}_i(t,\mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}_q),\mathbf{u})\phi_j(\boldsymbol{\xi}_q) = 0 \tag{2.30}$$

Integrating with respect to time t yields:

$$\sum_{q=1}^{M}(x_i(t,\boldsymbol{\xi}_q) - x_i(t_0,\boldsymbol{\xi}_q))\phi_j(\boldsymbol{\xi}_q)w_q - \int_{t_0}^{t}\sum_{q=1}^{M} w_q \mathbf{f}_i(t,\mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}_q),\mathbf{u})\phi_j(\boldsymbol{\xi}_q)dt = 0 \tag{2.31}$$

Interchanging the order of time integration and quadrature summation leads to

$$\sum_{q=1}^{M}\left\{x_i(t,\boldsymbol{\xi}_q) - x_i(t_0,\boldsymbol{\xi}_q) - \int_{t_0}^{t}\mathbf{f}_i(t,\mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}_q),\mathbf{u})dt\right\}\phi_j(\boldsymbol{\xi}_q)w_q = 0 \quad i = 1,\cdots,n \tag{2.32}$$

Note that the integral expression in Eq. (2.32) can be evaluated by an integration of the model equation with a specific instance of the random variable $\boldsymbol{\xi}_q$. Thus the process of evaluating the statistics on the output of the system reduces to sampling the chosen input points guided by quadrature method. Finally, the coefficients of the gPC expansion can be obtained as:

$$x_{i_k}(t) = \frac{1}{d_k^2}\sum_{q=1}^{M}\mathcal{X}_i(t_0,t,\boldsymbol{\xi}_q,\mathbf{u})\phi_k(\boldsymbol{\xi}_q)w_q, \qquad k,j = 0,1,\cdots,N, \quad i = 1,2,\cdots,n \tag{2.33}$$

where

$$\mathcal{X}_i(t_0,t,\boldsymbol{\xi}_q,\mathbf{u}) = x_i(t_0,\boldsymbol{\xi}_q) + \int_{t_0}^{t}\mathbf{f}_i(t,\mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{pc}\Phi(\boldsymbol{\xi}_q),\mathbf{u}) \tag{2.34}$$

$$d_k^2 = \int_{\Omega}\phi_k(\boldsymbol{\xi})\phi_k(\boldsymbol{\xi})p(\boldsymbol{\xi})d\boldsymbol{\xi} \tag{2.35}$$

Hence, the resulting method can be viewed as a MC-like evaluation of system equations, but with sample points selected by quadrature rules. PCQ approximates the moment of system state $\dot{\mathbf{x}} = f(t,\mathbf{x},\boldsymbol{\Theta},u)$ as:

$$\begin{aligned}\mathcal{E}[x_i(t)^N] &= \int_{\Omega}\left(\int_{t_0}^{t}\dot{x}_i dt\right)^N dp(\boldsymbol{\xi}) \\ &= \int_{\Omega}\left(x_i(t_0,\boldsymbol{\xi}) + \int_{t_0}^{t}\mathbf{f}_i(t,\mathbf{x},\boldsymbol{\Theta},\mathbf{u})dt\right)^N dp(\boldsymbol{\xi}) \quad i = 1,2,\cdots,n\end{aligned} \tag{2.36}$$

For a fixed value of parameter $\mathbf{\Theta} = \mathbf{\Theta}_q$, the time integration can be performed using deterministic integration. Integration (by PCQ) over the uncertain inputs determines the state pdf. This yields moment evaluations

$$\mathcal{E}[x_i(t)^N] = \sum_q w_q \left[\mathcal{X}_i(t_0, t, \boldsymbol{\xi}_q, \mathbf{u})\right]^N \quad i = 1, 2, \cdots, n \tag{2.37}$$

Thus the output moments can be approximated as a weighted sum of the outputs of simulations run at selected values of the uncertain input parameters (the quadrature points). The natural choice for these quadrature points is the set of Gaussian quadrature points which is defined by choosing the points optimally in the sense of maximizing the degree of polynomial function that integrates exactly. The classic method of Gaussian quadrature exactly integrates polynomials up to degree $2N + 1$ with $N + 1$ quadrature points. The tensor product of 1-dimension quadrature points is used to generate quadrature points in general n-dimension parameter space. As a consequence of this, the number of quadrature points increases exponentially as number of input parameters increases. It should be noted that this PCQ approach can still suffer from under-integration error if insufficient number of samples are used. This necessitates the need for an adaptive or nested quadrature scheme to successively refine the accuracy by increasing the number of sample points such as Clenshaw-Curtis quadrature method [50], [51] for numerical integration.

2.3 Examples

After introduction and development of Polynomial Chaos theory, we show performance of gPC expansion theory for some uncertain dynamic systems by validating gPC solution with Monte Carlo solution of these systems. Three different examples have been considered for simulation purposes:

1. First order forced dynamic equation
2. Duffing Oscillator
3. Hovering Helicopter Model

2.3.1 First Order System

Let us consider a simple single state system:

$$\dot{x} + Kx = 2e^{-t/10}sin(2t), \quad x(0) = 0 \tag{2.38}$$

where K is assumed to be a uniformly distributed parameter over the interval $[0.5, 1.5]$. The analytical solution of Eq. (2.38) is:

$$x(t, K) = e^{-Kt}(\frac{4}{(K-0.1)^2+4} - \frac{2e^{Kt-t/10}[2cos(2t) - (K-0.1)sin(2t)]}{(K-0.1)^2+4}) \tag{2.39}$$

Fig. 2.2 shows convergence of the first three central moments of x at $t = 2$ *sec.* as a function of different number of Monte Carlo runs. From these plots, it is clear that one needs a minimal of 3×10^4 random samples for the convergence in first three moments of state x. Now according to the gPC methodology, Legendre polynomials are used for the gPC expansion of both $x(t)$ and K. Using the procedure outlined in Section 2.1, Eq. (2.38) can be converted into the following deterministic form:

$$\mathcal{M}\dot{\mathbf{X}}_{pc}(t) + \mathcal{K}\mathbf{X}_{pc} = \begin{pmatrix} 2e^{-t/10}sin(2t) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{2.40}$$

where,

$$\mathcal{M}_{i+1,j+1} = \langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle = \frac{1}{2i+1}\delta_{ij}, \quad i, j = 0, 1, \cdots, N \tag{2.41}$$

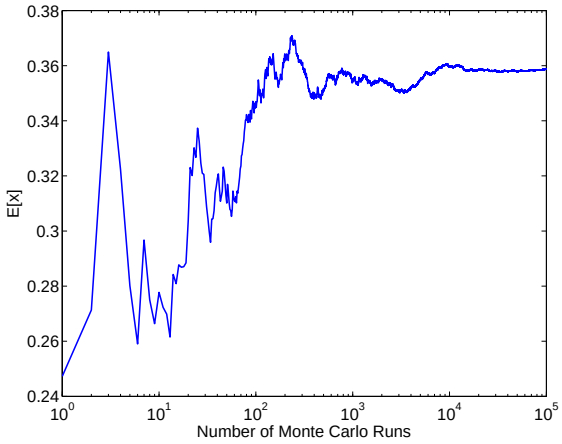
$$\mathcal{K}_{i+1,j+1} = \langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle + 0.5\langle \phi_1(\boldsymbol{\xi})\phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle, \quad i, j = 0, 1, \cdots, N \tag{2.42}$$

where, $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$, otherwise. $\mathcal{K}$ can be simplified as the following:

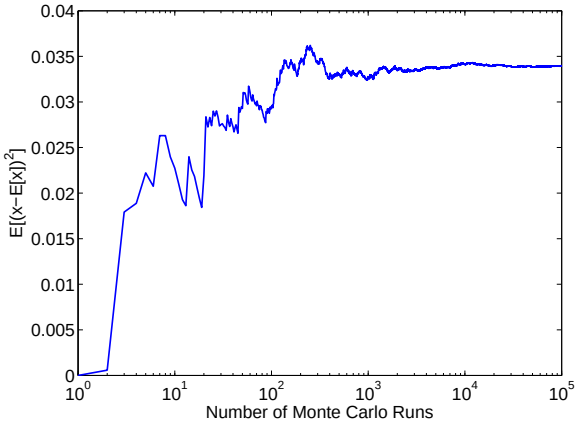
$$\mathcal{K} = \begin{cases} \frac{1}{2i+1}, & i = j \\ \frac{i}{(2i+1)(2i+3)}, & j = i+1 \\ \frac{i}{(2i-1)(2i+1)}, & j = i-1 \end{cases}$$

As well, the initial condition of Eq. (2.40) is given by

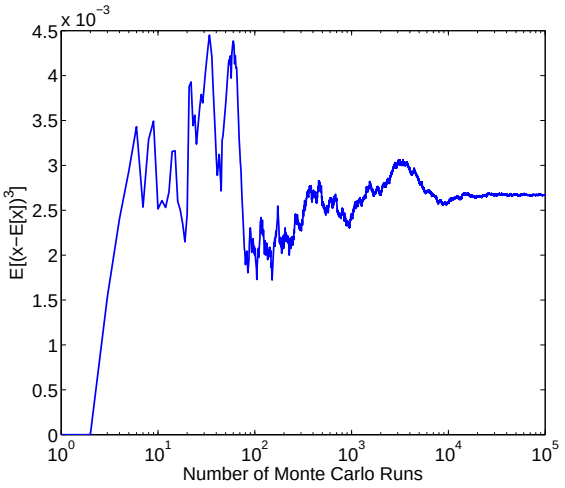
$$x_i(0) = 0 \quad i = 0, \cdots, N \tag{2.43}$$



(a) Mean

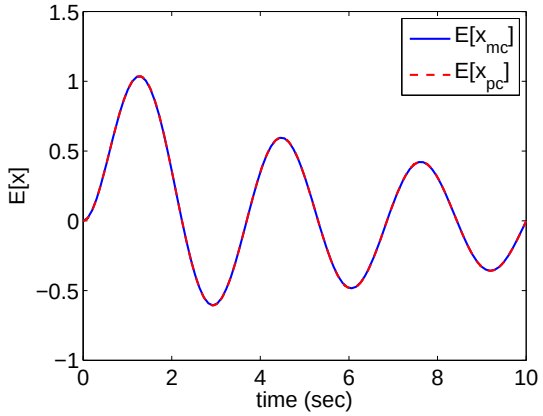


(b) Variance

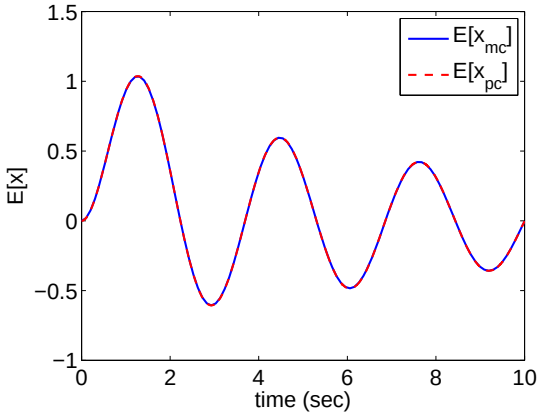


(c) 3^{rd} Central Moment

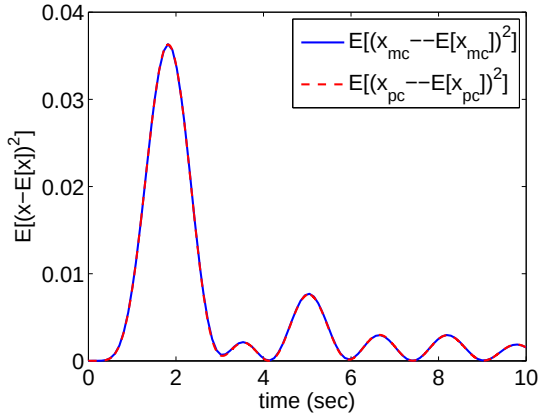
Figure 2.2: Convergence of Monte Carlo Solution at $t = 2$ *sec.* for Example 1.



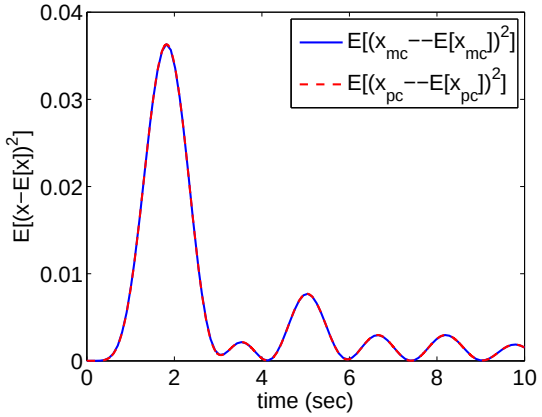
(a) Mean ($N = 2$)



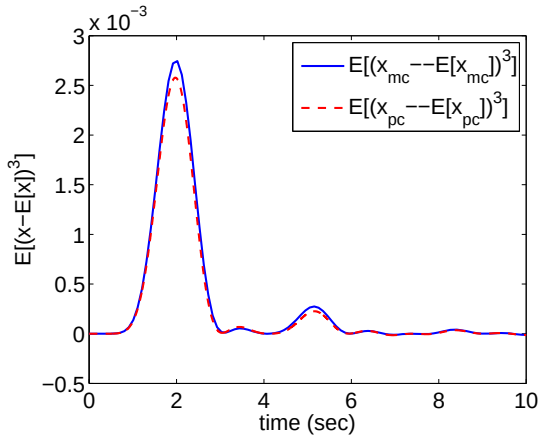
(b) Mean ($N = 4$)



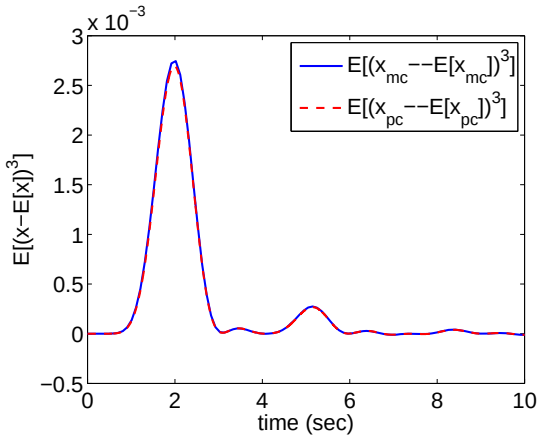
(c) Variance ($N = 2$)



(d) Variance ($N = 4$)

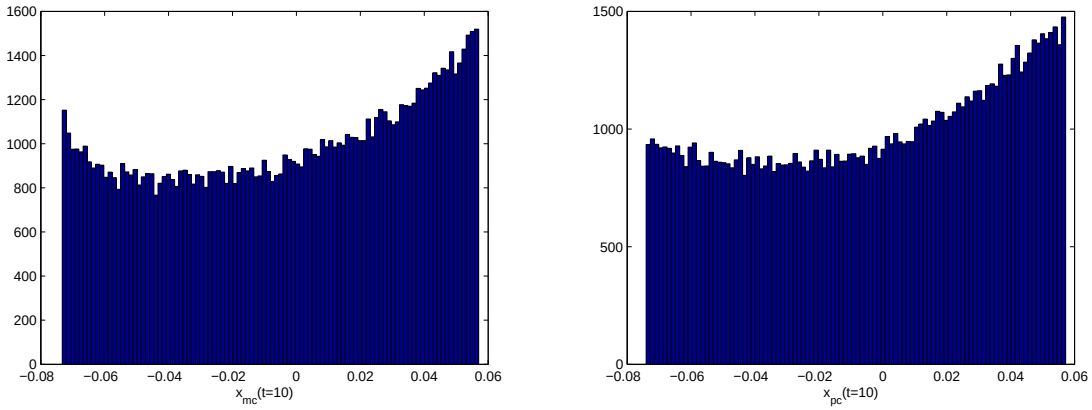


(e) 3^{rd} Central Moment ($N = 2$)



(f) 3^{rd} Central Moment ($N = 4$)

Figure 2.3: First Three Central Moments of x for Example 1.



(a) Monte Carlo

(b) 4^{th} order Polynomial Chaos expansion

Figure 2.4: Histogram of state x at $t = 10$ *sec.* for Example 1.

where, N is the number of terms used in the gPC expansion of x. Solution of this system of ODEs yields the coefficients of gPC expansion of $x(t)$, which can be used in Eq. (2.20) to construct the solution of Eq. (2.38).

Fig. 2.3 shows the evolution of first three central moments over time for different order gPC expansion. For comparison sake, we consider 10^5 MC runs to be the reference truth. From these plots, it is clear that first two moments are captured with a good accuracy with second order gPC expansion, i.e., $N = 2$, however, one needs at least fourth order gPC expansion to capture first three central moments. Furthermore, Fig. 2.4 shows histogram of state x at final time $t = 10$ *sec.* by using both 10^5 MC runs and 4^{th} order gPC solution. It is clear that both the gPC and Monte Carlo methods lead to similar distribution for x at final time.

Finally, Table 2.2 shows the relative error in approximating first three central moments using the PCQ framework for x at $t = 2$ seconds by assuming 10^5 MC runs to be reference truth. It should be noted that one needs only 4 quadrature points or model runs according to the PCQ formulation to capture first three moments with less than 1% error while 10^3 MC runs results in an order of magnitude higher error when compared against 10^5 MC runs. These results clearly show the efficacy of the gPC framework in accurately propagating the parameter uncertainty through dynamical system.

Table 2.2: Relative error of moments of state x with respect to 10^5 Monte Carlo runs at $t = 2sec.$

Number of Quadrature Points	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	9.10%	100%	100%
2	0.0617%	6.0777%	100%
3	0.0300%	0.0472%	5.0230%
4	0.0304%	0.0557%	0.0883%
10^3 MC Simulations	1.2310%	3.7051%	7.8948%

2.3.2 Duffing Oscillator

As the second example, let us consider the following nonlinear oscillator:

$$\ddot{x} + \eta\dot{x} + \alpha x + \beta x^3 = U_{in} \tag{2.44}$$

Eq. (2.44) represents a force input driven duffing oscillator with a cubic spring and a linear damping. For simulation purposes, we use $U_{in} = sin(3t)$ as force input function and $\beta = 2$, which is deterministic and α and η are considered to be uniformly distributed uncertain parameters over the intervals $[0.9, 1.4]$ and $[-1.45, -0.95]$, respectively. Also, initial conditions are assumed to be:

$$x(0) = -1, \qquad \dot{x}(0) = -1$$

The gPC expansion of initial distribution of α, η, $x(0)$ and $\dot{x}(0)$ can be written as:

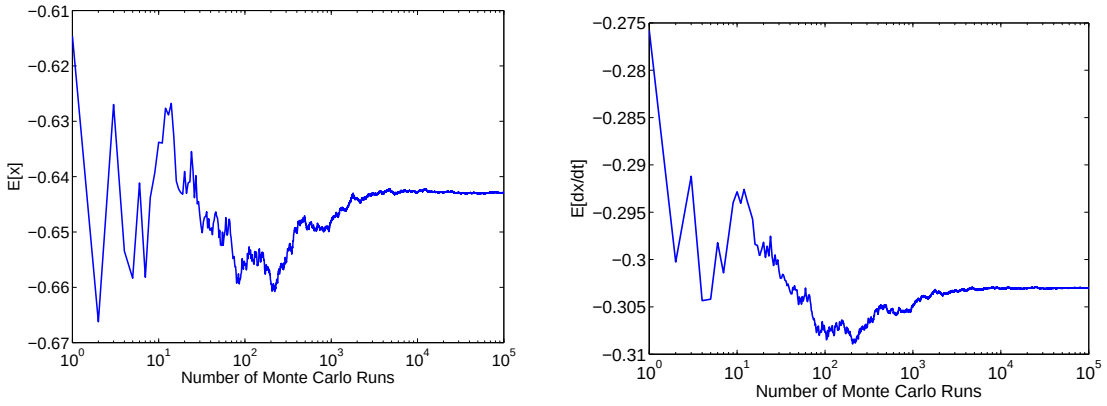
$$x(0, \boldsymbol{\xi}) = \sum_{k=0}^{N} x_k(0)\phi_k(\boldsymbol{\xi}) \qquad x_0(0) = -1 \quad x_k(0) = 0 \quad for \quad k \geq 2 \tag{2.45}$$

$$\dot{x}(0, \boldsymbol{\xi}) = \sum_{k=0}^{N} \dot{x}_k(0)\phi_k(\boldsymbol{\xi}) \qquad \dot{x}_0(0) = -1 \quad \dot{x}_k(0) = 0 \quad for \quad k \geq 2 \tag{2.46}$$

$$\eta(\boldsymbol{\xi}) = \sum_{k=0}^{N} \theta_{1_k}\phi_k(\boldsymbol{\xi}) \quad \theta_{1_0} = 1.15 \quad \theta_{1_1} = 0.25 \quad \theta_{1_k} = 0 \quad for \quad k > 2 \tag{2.47}$$

$$\alpha(\boldsymbol{\xi}) = \sum_{k=0}^{N} \theta_{2_k}\phi_k(\boldsymbol{\xi}) \quad \theta_{2_0} = -1.2 \quad \theta_{2_1} = 0.25 \quad \theta_{2_k} = 0 \quad for \quad k > 2 \tag{2.48}$$

where, ϕ_k are Legendre polynomials according to the gPC procedure. In this example, 6^{th} order polynomial chaos ($N = 6$) has been used to solve Eq. (2.44) and simulation time



(a) $\mathcal{E}[x]$ (b) $\mathcal{E}[\dot{x}]$

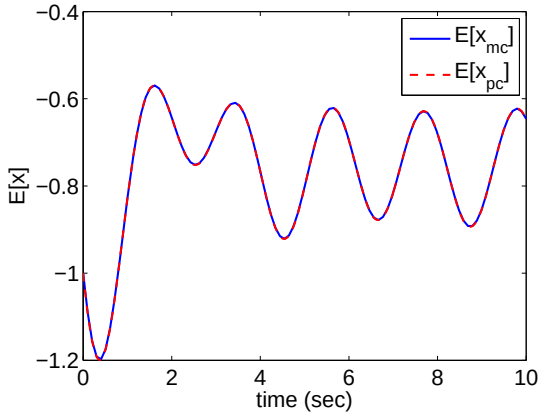
Figure 2.5: Moments of Monte Carlo solution at $t = 2$ *sec*.

interval is up to 10 seconds. Also, we have used 5 quadrature points in each direction of random variable $\boldsymbol{\xi}$ to evaluate Galerkin projection equations.

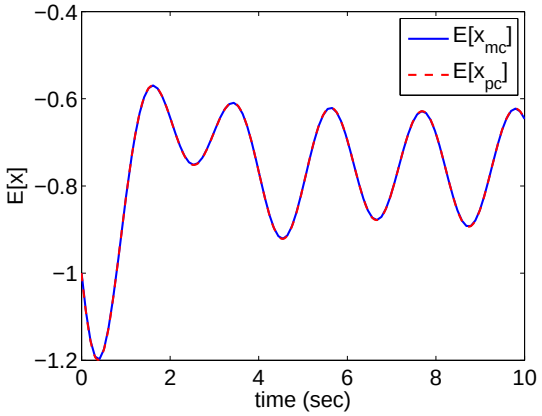
Convergence of the mean of states x and $\dot{x}$ has been shown in Fig. 2.5 as a function of different number of MC runs at $t = 2$ *sec*. From these plots, it is clear that one needs minimal 3×10^4 runs to guarantee convergence in the mean for both x and $\dot{x}$.

The gPC approximated first three central moments for state x and $\dot{x}$ are compared against those evaluated by using 10^5 Monte Carlo runs in Fig. 2.6 and Fig. 2.7, respectively. It is clear that the second order ($N = 2$) gPC expansion is able to capture the mean and variance for both x and $\dot{x}$, however, one needs 6^{th} order expansion to capture first three central moments. Furthermore, Tables 2.3 and 2.4 show the relative error in computing first three central moments for x and $\dot{x}$ using the PCQ framework and assuming 10^5 MC runs to be the reference truth. It is clear that one can obtain a better approximation for three central moments using only 9 quadrature points than 10^3 MC runs. It should be noted that the PCQ approximation error is within the convergence error of MC runs.

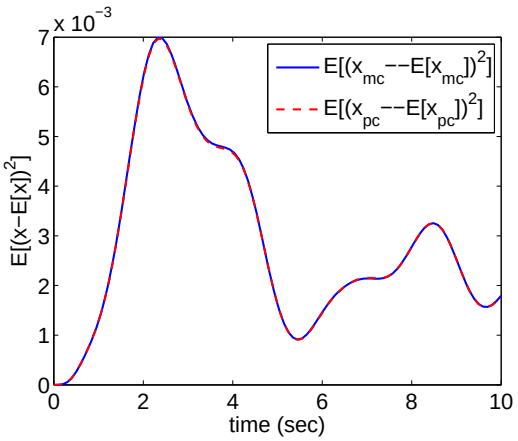
Finally, Fig. 2.8 and Fig. 2.9 shows the histograms for state x and $\dot{x}$ using 10^5 MC runs and 6^{th} order gPC expansion at $t = 10$ seconds. From these results, it is clear that the gPC expansion and MC method lead to similar distribution $t = 10$ *sec*.



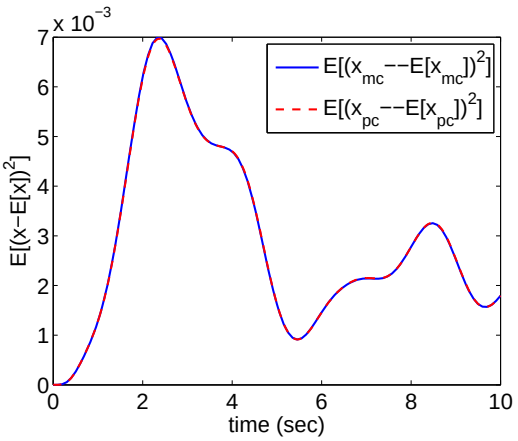
(a) Mean ($N = 2$)



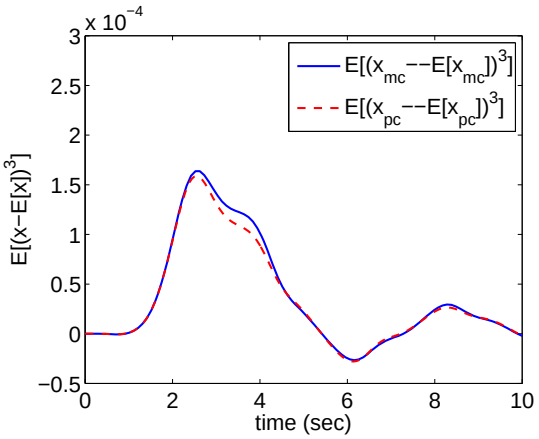
(b) Mean ($N = 6$)



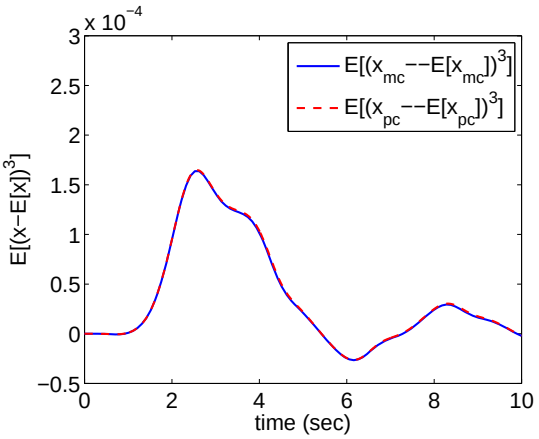
(c) Variance ($N = 2$)



(d) Variance ($N = 6$)

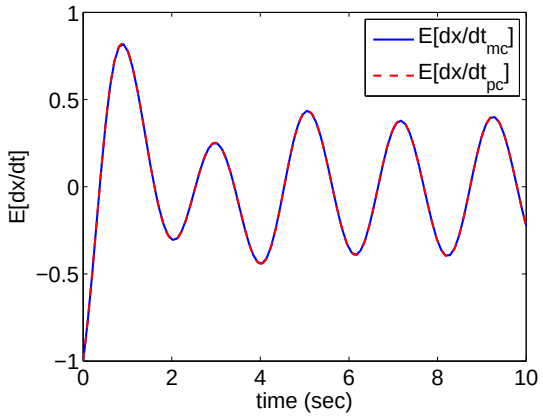


(e) 3^{rd} Central Moment ($N = 2$)

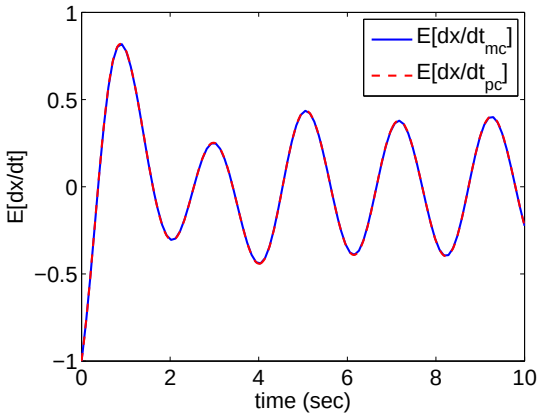


(f) 3^{rd} Central Moment ($N = 6$)

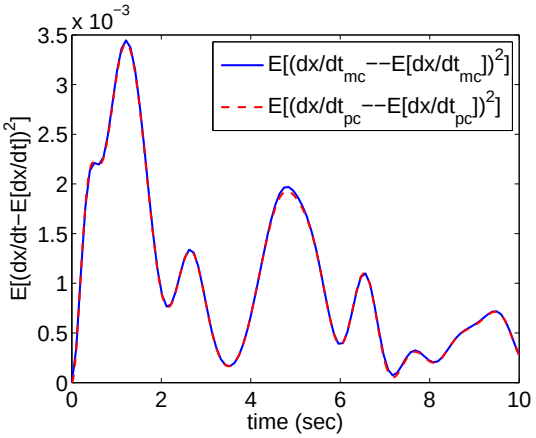
Figure 2.6: The first three central moments for x for Example 2.



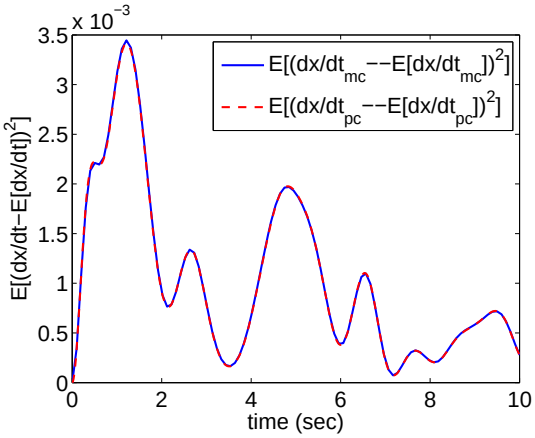
(a) Mean ($N = 2$)



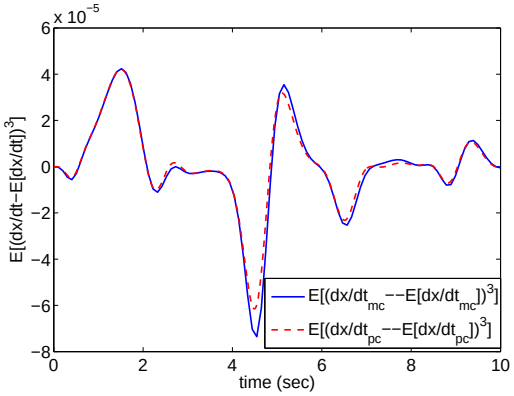
(b) Mean ($N = 6$)



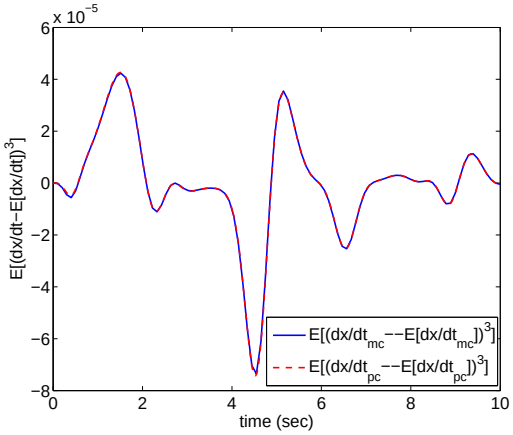
(c) Variance ($N = 2$)



(d) Variance ($N = 6$)



(e) 3^{rd} Central Moment ($N = 2$)



(f) 3^{rd} Central Moment ($N = 6$)

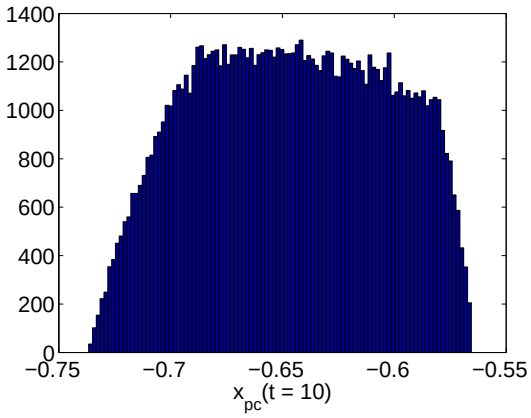
Figure 2.7: The first three central moments for $\dot{x}$ for Example 2.

Table 2.3: Relative error of moments of state x with respect to 10^5 Monte Carlo runs at $t = 2sec.$

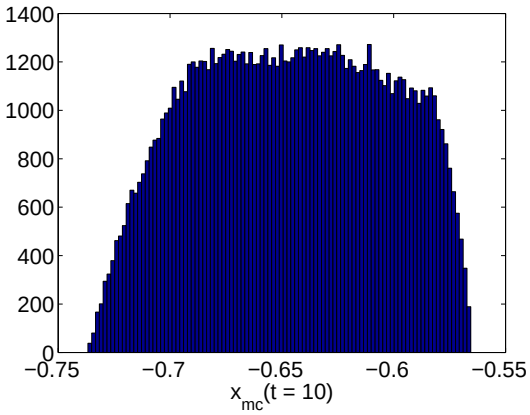
Number of Quadrature Points	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1^2	0.6545%	100%	100%
2^2	0.0279%	0.0239%	20.4861%
3^2	0.0269%	0.2522%	3.1240%
4^2	0.0269%	0.2537%	3.2633%
5^2	0.0269%	0.2537%	3.2642%
10^3 MC Simulations	0.8089%	0.5378%	19.9228%

Table 2.4: Relative error of moments of state $\dot{x}$ with respect to 10^5 Monte Carlo runs at $t = 2sec.$

Number of Quadrature Points	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1^2	2.0290%	100%	100%
2^2	0.0337%	1.2359%	63.9561%
3^2	0.0274%	0.1686%	2.3497%
4^2	0.0273%	0.2046%	2.9692%
5^2	0.0273%	0.2048%	3.0014%
10^3 MC Simulations	0.6299%	0.0791%	4.7318%

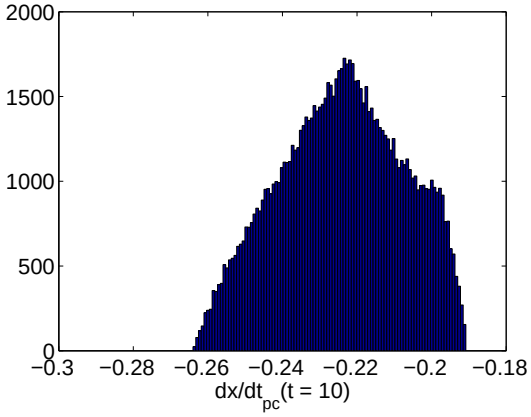


(a) 6^{th} order Polynomial Chaos expansion

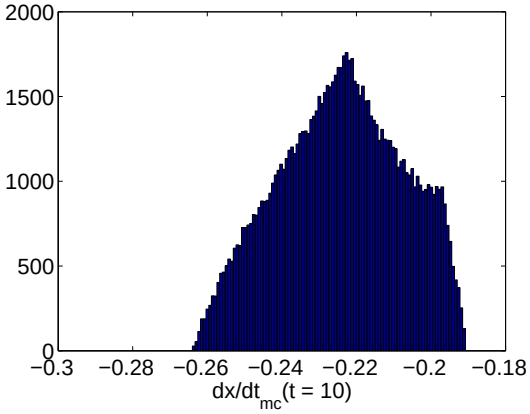


(b) Monte Carlo

Figure 2.8: Histogram of state x at $t = 10$ *sec.*



(a) 6^{th} order Polynomial Chaos expansion



(b) Monte Carlo

Figure 2.9: Histogram of state $\dot{x}$ at $t = 10$ *sec.* for Example 2.

2.3.3 Hovering Helicopter Model

As the last example, we examine efficiency of gPC method on a helicopter model [52] with the following system dynamics:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & -g & 0 \\ 1.26 & -1.765 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} - \begin{pmatrix} 0.086 \\ -7.408 \\ 0 \\ 0 \end{pmatrix} K_{lqr} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \tag{2.49}$$

where K_{lqr} and initial conditions are equal to:

$$K_{lqr} = [1.989 \quad -0.256 \quad -0.7589 \quad 1], \qquad X_{in} = [0.7929 \quad -0.0466 \quad -0.1871 \quad 0.5780]^T$$

Eq. (2.49) represents a helicopter model with an LQR controller, and actual value of p_1 and p_2 are $p_{1_{act}} = -0.0257$ and $p_{2_{act}} = 0.013$. We assume that p_1 and p_2 are uniformly distributed uncertain parameters over the intervals [-0.2, 0] and [0, 0.2], respectively. Similar to the previous example, initial conditions of states and initial distributions of uncertain parameters can be represented using the gPC expansion as:

$$x_i(0, \boldsymbol{\xi}) = \sum_{k=0}^{N} x_{i_k}(0)\phi_k(\boldsymbol{\xi}) \quad x_{i_k}(0) = 0 \ \text{ for } k = 1, 2, \cdots, N \text{ and } i = 1, \cdots, 4 \tag{2.50}$$

where,

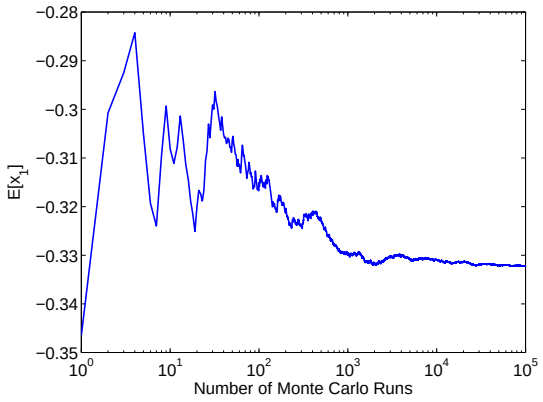
$$x_{1_0}(0) = 0.7929, \quad x_{2_0}(0) = -0.0466, \quad x_{3_0}(0) = -0.1871, \quad x_{4_0}(0) = 0.5780$$

and

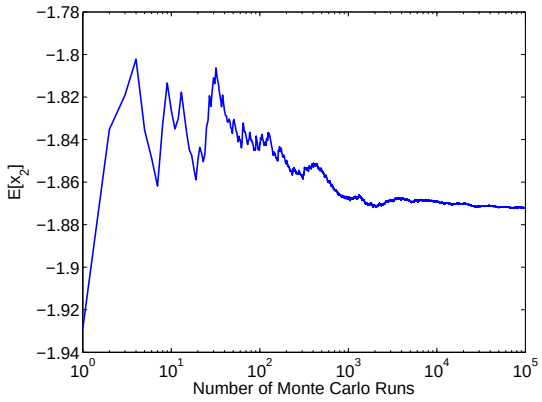
$$p_1(\boldsymbol{\xi}) = \sum_{k=0}^{N} a_{1_k}\phi_k(\boldsymbol{\xi}) \quad \theta_{1_0} = -0.1, \quad \theta_{1_1} = 0.1 \quad and\ \theta_{1_k} = 0 \ \ for \ \ k = 2, 3, \cdots, N \tag{2.51}$$

$$p_2(\boldsymbol{\xi}) = \sum_{k=0}^{N} a_{2_k}\phi_k(\boldsymbol{\xi}) \quad \theta_{2_0} = 0.1, \quad \theta_{2_1} = 0.1 \quad and\ \theta_{2_k} = 0 \ \ for \ \ k = 2, 3, \cdots, N \tag{2.52}$$

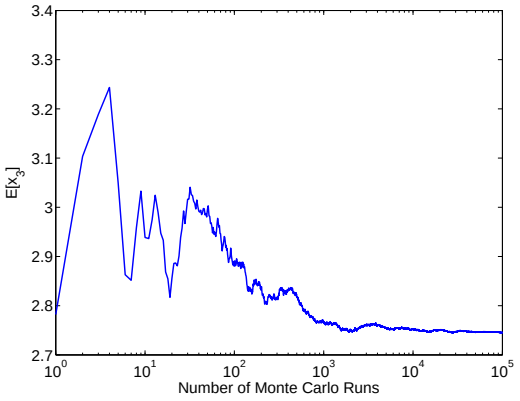
Fig. 2.10 shows convergence of the mean of states x_i $(i = 1, \cdots, 4)$ as a function of different number of Monte Carlo runs at $t = 2$ *sec.* It is clear that one needs at least 4×10^4 MC runs to get convergence in mean for all states. Furthermore, Tables 2.5 shows the relative error in computing first three central moments for x_1 using the PCQ framework



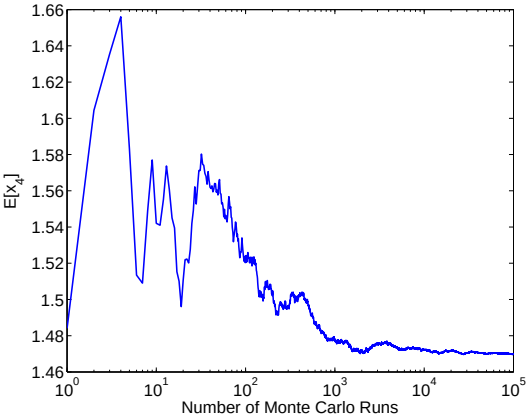
(a) $E(x_1)$



(b) $E(x_2)$

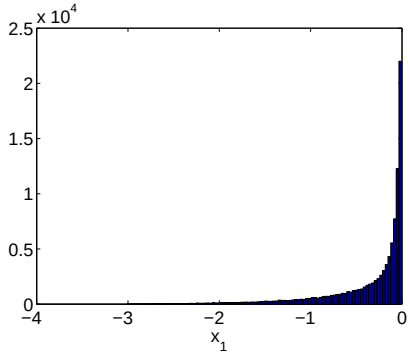


(c) $E(x_3)$

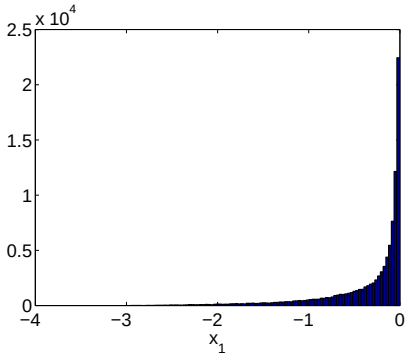


(d) $E(x_4)$

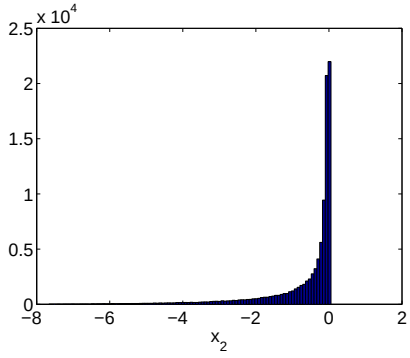
Figure 2.10: Convergence of Mean of Monte Carlo solutions at $t = 2$ *sec.* for Example 3.



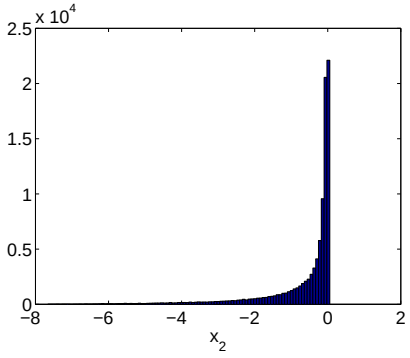
(a) x_1 (gPC approximation)



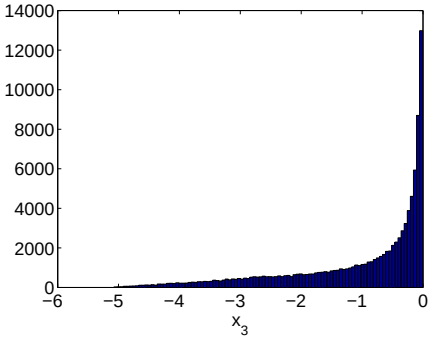
(b) x_1 (MC approximation)



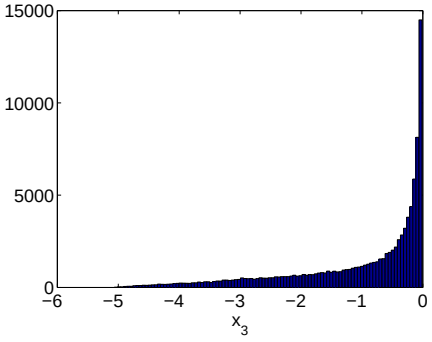
(c) x_2 (gPC approximation)



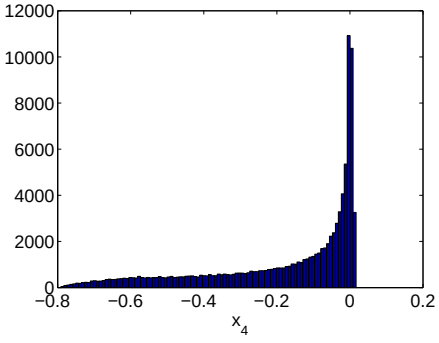
(d) x_2 (MC approximation)



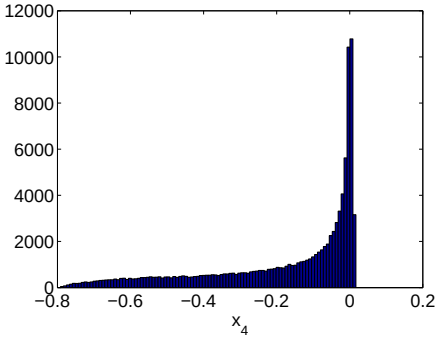
(e) x_3 (gPC approximation)



(f) x_3 (MC approximation)



(g) x_4 (gPC approximation)



(h) x_4 (MC approximation)

Figure 2.11: Histogram of states at $t = 10$ *sec.* for Example 3.

Table 2.5: Relative error of moments of state x_1 with respect to 10^5 Monte Carlo runs at $t = 2sec.$

Number of Quadrature Points	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1^2	0.1200%	100%	100%
2^2	0.03%	0.16%	11.00%
3^2	0.03%	0.5839%	0.8871%
10^3 MC Simulations	40.4619%	5.5996%	5.3558%

and assuming 10^5 MC runs to be the reference truth. It is clear that one can obtain a better approximation for three central moments using only 9 quadrature points than 10^3 MC runs. It should be noted that the PCQ approximation error is within the convergence error of MC runs. Finally, Fig. 2.11 shows the histogram plots for all states at $t = 10$ seconds using the 6^{th} order gPC expansion and 10^5 MC runs. From these plots, it is clear that both gPC and Monte Carlo methods lead to similar distributions for states.

2.4 Concluding Remarks

In this chapter, the main idea of the gPC expansion theory is discussed in detail and three numerical examples including both linear and nonlinear systems are considered to show the efficacy of the gPC methodology in capturing the non-Gaussian behavior. Ideally, infinite number of terms are required in the gPC expansion to capture the complete spectral content of the state pdf, i.e., all moments. However, the finite series truncation will always result in the error in capturing the state pdf. For examples considered in this paper, the gPC methodology is very competitive with the Monte Carlo approach in capturing higher order moments accurately with reasonable computational burden. In all of the considered examples, the gPC approach is much more numerically efficient than the Monte Carlo method.

Chapter 3

Estimation Process

In the previous chapter, the generalized Polynomial Chaos (gPC) theory is discussed in detail as a tool to propagate the state and parameter uncertainty through a nonlinear dynamic model. The use of sensor data to correct and refine the dynamical model forecast so as to reduce the associated uncertainty is a logical improvement over purely model-based prediction. However, mathematical models for various sensors are generally based upon the "usefulness" rather than the "truth" and do not provide all the information that one would like to know. Care must be taken when assimilating the observational data. As discussed in Chapter 1, there is currently no generic theoretical framework that solves the nonlinear filtering problem accurately and in a computationally efficient manner. Hence, there is a need to develop statistically and computationally efficient nonlinear filtering algorithms while appropriately accounting for the uncertainty in process and measurement models.

In this Chapter, two different gPC based approaches have been developed to design finite-dimension nonlinear filtering algorithms to integrate multiple sources of complementary information with system dynamics to help reduce the uncertainty of the output. Both the approaches make use of the gPC methodology to evaluate the high fidelity prediction between two measurement intervals. The first proposed method makes use of Bayes' formula to update gPC series expansion while the second method updates the gPC series expansion based upon minimum variance estimator.

3.1 Fusion of Measurement Data and Process Model

Given a prediction model of Eq. (2.19), let us assume the following sensor model to obtain the measurement data:

$$\mathbf{y}_k \triangleq \mathbf{y}(t_k) = \mathbf{h}(\mathbf{x}_k, \mathbf{\Theta}) + \boldsymbol{\nu}_k \tag{3.1}$$

where, $\mathbf{y}_k \in \mathbb{R}^k$ is observed sensor data and $\boldsymbol{\nu} \in \mathbb{R}^k$ is the measurement noise with prescribed likelihood function $p(\mathbf{y}_k|\mathbf{x}_k)$ which is generally assumed to be a zero mean Gaussian pdf. Using the gPC uncertainty evolution as a forecasting tool , the joint state and parameter pdf can be updated using the Bayes' rule on the arrival of a measurement data:

$$p(\mathbf{\Theta}, \mathbf{x}|\mathbf{Y}_k) = \frac{p(\mathbf{\Theta}, \mathbf{x}|\mathbf{Y}_{k-1})p(\mathbf{y}(t_k)|\mathbf{\Theta}, \mathbf{x})}{p(\mathbf{y}(t_k))} \tag{3.2}$$

where $\mathbf{Y}_k$ represents the measurement data up to time t_k. $p(\mathbf{\Theta}, \mathbf{x}|\mathbf{Y}_{k-1})$ is the joint prior pdf (solution of the gPC approach) of $\mathbf{x}$ and $\mathbf{\Theta}$ at time t_k given all observations up to time t_{k-1}, $p(\mathbf{y}_k|\mathbf{\Theta}, \mathbf{x})$ is the likelihood that we observe $\mathbf{y}_k$ given $\mathbf{x}$ and $\mathbf{\Theta}$ at time t_k and $p(\mathbf{\Theta}, \mathbf{x}|\mathbf{Y}_k)$ represents the joint posterior pdf of $\mathbf{x}$ and $\mathbf{\Theta}$ at time t_k given all previous observations, including $\mathbf{y}_k$. Furthermore, $p(\mathbf{y}_k)$ is the total probability of observation at time t_k which can be evaluated as follows:

$$p(\mathbf{y}_k) = \int\int p(\mathbf{\Theta}, \mathbf{x}|\mathbf{Y}_{k-1})p(\mathbf{y}_k|\mathbf{\Theta}, \mathbf{x})d\mathbf{\Theta}d\mathbf{x} \tag{3.3}$$

As we concluded in the previous chapter, the gPC approach provides us a tool to determine equations of evolutions for conditional moments for the prior joint pdf $p(\mathbf{\Theta}, \mathbf{x}|\mathbf{Y}_{k-1})$. We now seek to develop equations of evolutions for the posterior conditional moments. As a step towards this goal, let us consider a continuously differentiable scalar function $\phi(\mathbf{\Theta}, \mathbf{x})$ and define posterior and prior conditional moments as:

$$\hat{\phi}_k^+ = \mathcal{E}^+[\phi(\mathbf{\Theta}, \mathbf{x})] \triangleq \int\int \phi(\mathbf{\Theta}, \mathbf{x})p(\mathbf{\Theta}, \mathbf{x}|\mathbf{Y}_k)d\mathbf{\Theta}d\mathbf{x} \tag{3.4}$$

$$\hat{\phi}_k^- = \mathcal{E}^-[\phi(\mathbf{\Theta}, \mathbf{x})] \triangleq \int\int \phi(\mathbf{\Theta}, \mathbf{x})p(\mathbf{\Theta}, \mathbf{x}|\mathbf{Y}_{k-1})d\mathbf{\Theta}d\mathbf{x} \tag{3.5}$$

Now, multiplying Eq. (3.2) with $\phi(\mathbf{\Theta}, \mathbf{x})$ and integrating over $\mathbf{\Theta}$ and $\mathbf{x}$, we get:

$$\hat{\phi}_k^+ = \frac{\mathcal{E}^-[\phi(\mathbf{\Theta}, \mathbf{x})p(\mathbf{y}_k|\mathbf{\Theta}, \mathbf{x})]}{p(\mathbf{y}_k)} \tag{3.6}$$

Note that $\mathbf{y}_k$ is fixed with respect to the expectation operator and thus, the right-hand side of Eq. (3.6) is a function of $\mathbf{y}_k$ only. Notice that Eq. (3.6) is not an ordinary difference equation and the evaluation of right-hand side of Eq. (3.6) requires the knowledge of the prior density function. Thus, even the computation of the posterior mean for $\boldsymbol{\Theta}$ and $\mathbf{x}$, i.e., $\phi = \boldsymbol{\Theta}$ or $\mathbf{x}$ depends upon all the other moments. In the next section, we shall present the details to obtain a computationally realizable filter in the general nonlinear case while making use of the gPC expansion series. For the sake of simplicity, we shall assume the likelihood function to be a normal density function although the development present in next section is applicable to any generic likelihood function.

$$\begin{aligned} p(\mathbf{y}_k|\boldsymbol{\Theta},\mathbf{x}) &= \mathcal{N}(\mathbf{y}_k|\mathbf{h}(\mathbf{x}(t),\boldsymbol{\Theta}),\mathbf{R}_k) \\ &\triangleq \frac{1}{\sqrt{(2\pi)^k|R_k|}} e^{-\frac{1}{2}\left(\mathbf{y}_k-\mathbf{h}(\mathbf{x}(t),\boldsymbol{\Theta})\right)^T\mathbf{R}_k^{-1}\left(\mathbf{y}_k-\mathbf{h}(\mathbf{x}(t),\boldsymbol{\Theta})\right)} \end{aligned} \tag{3.7}$$

3.2 gPC-Bayes Approach

As discussed in the last section, the main challenge during the measurement update process lies in evaluating expectation integrals involved in Eq. (3.6) in a numerically efficient way. Although the gPC process does not provide us a closed-form expression for the state or parameter pdf but it can be used effectively in computing the expectation integrals. As discussed in the previous chapter, all moments of random variables $\boldsymbol{\Theta}$ and $\mathbf{x}$ are just function of their gPC expansion coefficients, i.e., $\boldsymbol{\Theta}_{pc}$ and $\mathbf{X}_{pc}$. Hence, one can just update the gPC coefficients on the arrival of measurement data based upon Eq. (3.6). So if we define $\boldsymbol{\Theta}_{pc}^-$ and $\mathbf{X}_{pc}^-$ to be the prior gPC coefficients and $\boldsymbol{\Theta}_{pc}^+$ and $\mathbf{X}_{pc}^+$ to be posterior gPC coefficients, then we can evaluate $\hat{\phi}^-(\boldsymbol{\Theta},\mathbf{x})$ and $\hat{\phi}^+(\boldsymbol{\Theta},\mathbf{x})$ as:

$$\hat{\phi}_k^- = \hat{\phi}^-(\boldsymbol{\Theta},\mathbf{x}) = \mathcal{E}^-[\phi(\boldsymbol{\Theta},\mathbf{x})] = \int \phi(\boldsymbol{\Theta}_{pc}^-\Phi(\boldsymbol{\xi}),\mathbf{X}_{pc}^-(t)\Phi(\boldsymbol{\xi}))p(\boldsymbol{\xi})d\boldsymbol{\xi} \tag{3.8}$$

$$\hat{\phi}_k^+ = \hat{\phi}^+(\boldsymbol{\Theta},\mathbf{x}) = \mathcal{E}^+[\phi(\boldsymbol{\Theta},\mathbf{x})] = \int \phi(\boldsymbol{\Theta}_{pc}^+\Phi(\boldsymbol{\xi}),\mathbf{X}_{pc}^+(t)\Phi(\boldsymbol{\xi}))p(\boldsymbol{\xi})d\boldsymbol{\xi} \tag{3.9}$$

Similarly, the $\mathcal{E}^-[\phi(\boldsymbol{\Theta},\mathbf{x})p(\mathbf{y}_k|\boldsymbol{\Theta},\mathbf{x})]$ can be evaluated as:

$$\begin{aligned} M_r(\boldsymbol{\Theta}_{pc}^-,\mathbf{x}_{pc}^-,\mathbf{y}_k) &= \mathcal{E}^-[\phi(\boldsymbol{\Theta},\mathbf{x})p(\mathbf{y}_k|\boldsymbol{\Theta},\mathbf{x})] \\ &= \int \underbrace{\phi(\boldsymbol{\Theta}_{pc}^-\Phi(\boldsymbol{\xi}),\mathbf{X}_{pc}^-(t)\Phi(\boldsymbol{\xi}))\mathcal{N}(\mathbf{y}_k|\mathbf{h}(\mathbf{X}_{pc}^-(t)\Phi(\boldsymbol{\xi}),\boldsymbol{\Theta}_{pc}^-\Phi(\boldsymbol{\xi})),\mathbf{R}_k)}_{\psi(\boldsymbol{\Theta}_{pc}^-\Phi(\boldsymbol{\xi}),\mathbf{X}_{pc}^-(t)\Phi(\boldsymbol{\xi}),\mathbf{y}_k,\mathbf{R}_k)} p(\boldsymbol{\xi})d\boldsymbol{\xi} \end{aligned} \tag{3.10}$$

For moment evaluation purpose, $\phi(\mathbf{\Theta}, \mathbf{X})$ is a polynomial function and one can obtain a closed-form expressions for $\hat{\phi}_k^-$ and $\hat{\phi}_k^+$. For example, the posterior mean and covariance are given as:

$$\mathcal{E}[x_i^+(t)] = x_{i_1}(t), \quad i = 1, \cdots, n \tag{3.11}$$

$$\mathcal{E}[x_i^+(t)x_j^+(t)] = \sum_{k=0}^{N} x_{i_k}^+(t)x_{j_k}^+(t), \quad i, j = 1, \cdots, n \tag{3.12}$$

The main challenge lies in evaluating $M_r(\mathbf{\Theta}_{pc}^-, \mathbf{x}_{pc}^-, \mathbf{y}_k)$. As discussed in the last Chapter, one can use quadrature scheme to evaluate $M_r(\mathbf{\Theta}_{pc}^-, \mathbf{x}_{pc}^-, \mathbf{y}_k)$:

$$\begin{aligned} M_r(\mathbf{\Theta}_{pc}^-, \mathbf{x}_{pc}^-, \mathbf{y}_k) &\approx \sum_{q=1}^{N_q} w_q \psi(\mathbf{\Theta}_{pc}^- \Phi(\boldsymbol{\xi}_q), \mathbf{X}_{pc}^-(t)\Phi(\boldsymbol{\xi}_q), \mathbf{y}_k, \mathbf{R}_k) \\ &= \sum_{q=1}^{N_q} w_q \psi(\mathbf{\Theta}_q^-, \mathbf{x}_q^-, \mathbf{y}_k, \mathbf{R}_k) \end{aligned} \tag{3.13}$$

Notice that $M_r(\mathbf{\Theta}_{pc}^-, \mathbf{x}_{pc}^-, \mathbf{y}_k)$ is completely known since prior values of coefficients are known from the gPC solution of the system. Also $\phi(\mathbf{\Theta}, \mathbf{x})$ takes the following form to match all joint moments up to order N_m:

$$\phi(\mathbf{\Theta}, \mathbf{x}) = \mathbf{\Theta}_i^{s_1} \mathbf{x}_j^{s_2}, \ s_1 + s_2 \leq N_m \tag{3.14}$$

Now, substitution of Eq. (3.9) and Eq. (3.13) and in Eq. (3.6) leads to N_c nonlinear coupled equations which defines posterior gPC coefficients $\mathbf{\Theta}^+$ and $\mathbf{x}^+$ in terms of prior information which is available from measurement and gPC propagation to match all joint moments up to order N_m:

$$g_{s_1,s_2}(\mathbf{\Theta}_{pc}^+, \mathbf{X}_{pc}^+) = \hat{\phi}^+(\mathbf{\Theta}_{pc}^+, \mathbf{x}_{pc}^+) - \frac{1}{\alpha} M_r(\mathbf{\Theta}_{pc}^-, \mathbf{x}_{pc}^-, \mathbf{y}_k), \ \alpha = p(\mathbf{y}_k), \ s_1 + s_2 \leq N_m \tag{3.15}$$

If we define $f_i(\mathbf{\Theta}_{pc}^+, \mathbf{X}_{pc}^+) = g_{s_1,s_2}(\mathbf{\Theta}_{pc}^+, \mathbf{X}_{pc}^+)$, then one can pose the following minimization problem to find a solution for posterior coefficients $\mathbf{\Theta}_{pc}^+$ and $\mathbf{X}_{pc}^+$.

$$\min_{\mathbf{\Theta}_{pc}^+, \mathbf{X}_{pc}^+} \left(\sum_{i=1}^{N_c} f_i^2(\mathbf{\Theta}_{pc}^+, \mathbf{X}_{pc}^+) \right) \tag{3.16}$$

where N_c is given as:

$$N_c = \sum_{k=1}^{N_m} \frac{(m+n)!}{k!(m+n-k)!} \frac{(N_m)!}{k!(N_m-k)!} \tag{3.17}$$

and n and m are dimension of state $\mathbf{x}$ and parameter $\boldsymbol{\Theta}$, respectively. Different algorithms like Trust-Region-Reflective Optimization [53], [54], Levenberg-Marquardt Optimization [55], [56], [57] and Gauss-Newton approach [57], [58] can be used to solve this optimization problem. In this thesis, we have used Levenberg-Marquardt to solve this optimization problem. For a special case for matching just posterior mean, i.e., $N_m = 1$, we get the following analytical solution for the posterior coefficients:

$$\boldsymbol{\Theta}^{+}_{pc_1} = \kappa_{1,0} \tag{3.18}$$

$$\mathbf{X}^{+}_{pc_1} = \kappa_{0,1} \tag{3.19}$$

where, $\boldsymbol{\Theta}^{+}_{pc_1}$ and $\mathbf{X}^{+}_{pc_1}$ represents the first column of $\boldsymbol{\Theta}^{+}_{pc}$ and $\mathbf{X}^{+}_{pc}$, respectively. $\kappa_{1,0}$ and $\kappa_{0,1}$ are given as:

$$\kappa_{1,0} = \sum_{q=1}^{N_q} w_q \left[\boldsymbol{\Theta}^{-}_{pc}\Phi(\boldsymbol{\xi})\right] \mathcal{N}(\mathbf{y}_k|\mathbf{h}(\mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}_q), \boldsymbol{\Theta}^{-}_{pc}\Phi(\boldsymbol{\xi}_q)), \mathbf{R}_k) \tag{3.20}$$

$$\kappa_{0,1} = \sum_{q=1}^{N_q} w_q \left[\mathbf{X}^{-}_{pc}(t)\Phi(\boldsymbol{\xi})\right] \mathcal{N}(\mathbf{y}_k|\mathbf{h}(\mathbf{X}_{pc}(t)\Phi(\boldsymbol{\xi}_q), \boldsymbol{\Theta}^{-}_{pc}\Phi(\boldsymbol{\xi}_q)), \mathbf{R}_k) \tag{3.21}$$

Since the only moment constraint is expected value of states and parameters, the gPC-Bayes approach just updates coefficient of just the first term in the gPC expansion of state $\mathbf{x}$ and parameter $\boldsymbol{\Theta}$, and retains prior value of the rest of the coefficients.

3.3 Polynomial Chaos Based Minimum Variance Estimator

In the previous section, we developed an estimation algorithm to estimate posterior moments and gPC expansion coefficients by making use of the Bayes' rule. In this section, we present an alternative development based upon minimum variance estimator. The main advantage of the minimum variance approach is that it is computationally more efficient than the gPC-Bayes method and it is easier to implement.

3.3.1 Minimum Variance Estimation with a Priori Information

Let us consider augmented state vector $\mathbf{z}$ consisting of both state and parameters:

$$\mathbf{z}(t,\boldsymbol{\xi}) = \begin{bmatrix} \mathbf{x}(t,\boldsymbol{\xi}) \\ \boldsymbol{\Theta}(\boldsymbol{\xi}) \end{bmatrix} \tag{3.22}$$

We define the prior mean $\hat{\mathbf{z}}_k^-$ and prior covariance $\mathbf{P}_k^-$ as follows:

$$\hat{\mathbf{z}}_k^- \triangleq \mathcal{E}^-[\mathbf{z}_k] = \begin{bmatrix} \mathbf{X}_{pc_1}^-(t) \\ \boldsymbol{\Theta}_{pc_1}^- \end{bmatrix} \tag{3.23}$$

$$\mathbf{P}_k{}^- \triangleq \mathcal{E}^-[(\mathbf{z}_k - \hat{\mathbf{z}}_k^-)(\mathbf{z}_k - \hat{\mathbf{z}}_k^-)^T] = \begin{pmatrix} \sum_{i=1}^{N} \mathbf{X}_{pc_i}^{-2} & \sum_{i=1}^{N} \mathbf{X}_{pc_i}^- \boldsymbol{\Theta}_{pc_i}^- \\ \sum_{i=1}^{N} \mathbf{X}_{pc_i}^- \boldsymbol{\Theta}_{pc_i}^- & \sum_{i=1}^{N} \boldsymbol{\Theta}_{pc_i}^{-2} \end{pmatrix} \tag{3.24}$$

where, $\mathbf{X}_{pc_i}^-$ and $\boldsymbol{\Theta}_{pc_i}^-$ are the i^{th} column of the gPC expansion coefficient matrices $\mathbf{X}_{pc}^-$ and $\boldsymbol{\Theta}_{pc}^-$, respectively. Given an estimate of prior mean and covariance, the posterior mean and covariance according to minimum variance formulation is given as [59]:

$$\hat{\mathbf{z}}_k^+ = \hat{\mathbf{z}}_k^- + \mathbf{K}_k[\mathbf{y}_k - \mathcal{E}^-[\mathbf{h}(\mathbf{x}_k, \boldsymbol{\Theta})]] \tag{3.25}$$

$$\mathbf{P}_k^+ = \mathbf{P}_k^- + \mathbf{K}_k \mathbf{P}_{zy} \tag{3.26}$$

$$\mathbf{K}_k = -\mathbf{P}_{zy}^T \left(\mathbf{P}_{hh}^- + \mathbf{R}_k\right)^{-1} \tag{3.27}$$

where, $\mathbf{K}_k$ is known as the Kalman gain matrix and matrices $\mathbf{P}_{zy}$ and $\mathbf{P}_{zz}$ are defined as:

$$\hat{\mathbf{h}}_k^- \triangleq \mathcal{E}^-[\mathbf{h}(\mathbf{x}_k, \boldsymbol{\Theta})] = \sum_{q=1}^{M} w_q \mathbf{h}(\mathbf{x}_k(\boldsymbol{\xi}_q), \boldsymbol{\theta}(\boldsymbol{\xi}_q)) \tag{3.28}$$

$$\mathbf{P}_{zy} \triangleq \mathcal{E}^-[(\mathbf{z}_k - \hat{\mathbf{z}}_k)(\mathbf{h}(\mathbf{x}_k, \boldsymbol{\Theta}) - \hat{\mathbf{h}}_k^-)^T] = \sum_{q=1}^{M} w_q (\mathbf{z}_k(\boldsymbol{\xi}_q) - \hat{\mathbf{z}}_k^-)(\underbrace{\mathbf{h}(\mathbf{x}_k(\boldsymbol{\xi}_q), \boldsymbol{\theta}(\boldsymbol{\xi}_q))}_{\mathbf{h}_q} - \hat{\mathbf{h}}_k^-)^T \tag{3.29}$$

$$\mathbf{P}_{hh}^- \triangleq \mathcal{E}^-[(\mathbf{h}(\mathbf{x}_k, \boldsymbol{\Theta}) - \hat{\mathbf{h}}_k^-)(\mathbf{h}(\mathbf{x}_k, \boldsymbol{\Theta}) - \hat{\mathbf{h}}_k^-)^T] = \sum_{q=1}^{M} w_q (\mathbf{h}_q - \hat{\mathbf{h}}_k^-)(\mathbf{h}_q - \hat{\mathbf{h}}_k^-)^T \tag{3.30}$$

Similar to prior mean and covariance, posterior mean $\hat{\mathbf{z}}_k^+$ and covariance $\mathbf{P}_k^+$ can also be written in terms of posterior gPC expansion coefficients for both state and parameters:

$$\hat{\mathbf{z}}_k^+ = \begin{bmatrix} \mathbf{X}_{pc_1}^+(t) \\ \boldsymbol{\Theta}_{pc_1}^+ \end{bmatrix} \tag{3.31}$$

$$\mathbf{P}_k^+ = \begin{pmatrix} \sum_{i=1}^{N} \mathbf{X}_{pc_i}^{+^2} & \sum_{i=1}^{N} \mathbf{X}_{pc_i}^+ \mathbf{\Theta}_{pc_i}^+ \\ \sum_{i=1}^{N} \mathbf{X}_{pc_i}^+ \mathbf{\Theta}_{pc_i}^+ & \sum_{i=1}^{N} \mathbf{\Theta}_{pc_i}^{+^2} \end{pmatrix} \tag{3.32}$$

Eq. (3.25) and Eq. (3.31) provide a closed-form solution for $\mathbf{X}_{pc_1^+}$ and $\mathbf{\Theta}_{pc_1}^+$ while one can solve for rest of the posterior coefficients while making use of Eq. (3.26) and Eq. (3.32).

Chapter 4

Numerical Simulation

In the previous chapter, we have developed two algorithms based upon the gPC expansion for state and parameter estimation. In this chapter, we consider four different numerical experiments to demonstrate performance of these methods. We also employ the EKF and bootstrap particle filter algorithms to compare the performance of proposed methodology.

4.1 First Example: First Order System

As the first example, we consider forced first order system mentioned in section 2.3.1:

$$\dot{x} + Kx = U_{in}, \quad x(0) = 0 \tag{4.1}$$

where, $U_{in} = 2e^{-t/10}sin(2t)$ and prior uncertainty in K is assumed to be uniformly distributed over the interval $[0.5, 1.5]$. For simulation purposes, measurement data is assumed to be available at sampling frequency of 1Hz. A random sample of K is taken from prior distribution to generate the true measurement data. The results presented in thesis corresponds to true value of K being 1.3659 ($K_{act} = 1.3659$). The true measurement data is corrupted with a Gaussian white noise of zero mean and variance being 0.05. To represent uncertainty in state and parameter, 9^{th} order gPC expansion is considered and total simulation time interval is assumed to be 10 *sec*. The initial gPC expansion for K and $x(0)$ can be written

as:

$$x(0,\boldsymbol{\xi}) = \sum_{k=0}^{9} x_k(0)\phi_k(\boldsymbol{\xi}) \quad x_k(0) = 0 \tag{4.2}$$

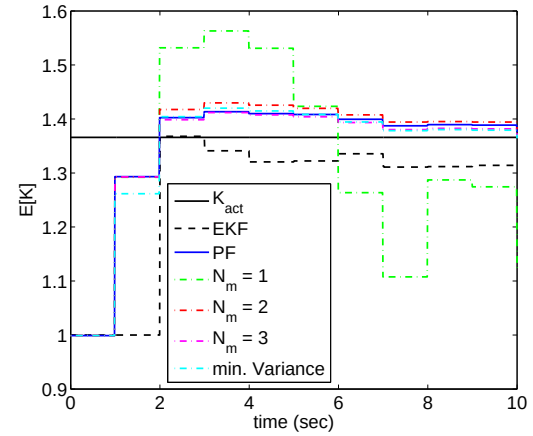
$$K(\boldsymbol{\xi}) = \sum_{i=0}^{9} k_i\phi_i(\boldsymbol{\xi}) \quad k_0 = 1, \quad k_2 = 0.5 \quad \text{and} \quad k_i = 0 \tag{4.3}$$

where, $\phi_k(\boldsymbol{\xi})'s$ are Legendre polynomial which correspond to uniform distribution of parameter K.

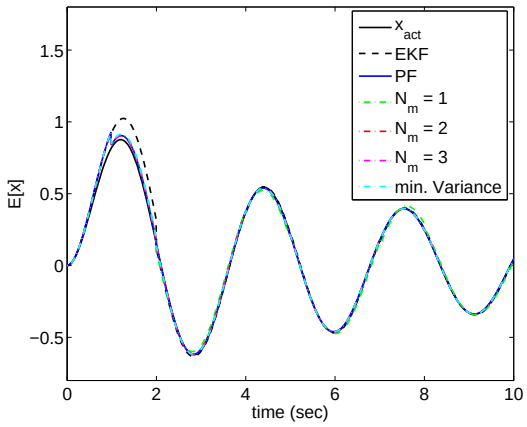
The mean estimates for parameter K and state x by using Particle Filter (PF), EKF, gPC based minimum variance estimator, and gPC-Bayes method for different moment matching constraints (different values of N_m) have been shown in Fig. 4.1(a) and Fig. 4.1(b), respectively. As expected, the gPC-Bayes method results in more accurate results as we increase N_m and assuming the PF approximated posterior mean to be the reference truth. Also, when $N_m = 2$, the gPC-Bayes and gPC based minimum variance estimator perform very similar in finding posterior mean estimates for both K and x. Both the EKF and the gPC-Bayes method with $N_m = 1$ perform poorly in the estimation of the first posterior moment of K and x.

Fig. 4.1(c) and Fig. 4.1(d) show posterior variance for parameter K and state x corresponding to different filters, respectively. As expected, the gPC-Bayes approach with $N_m = 1$ can not capture the posterior variance for parameter K and state x. However, the performance of the gPC-Bayes method improves a lot in capturing the posterior variance as compared to the PF estimates variance by increasing N_m, i.e., number of matching moment constraints. Once again, both the gPC-Bayes method and gPC based minimum variance estimator perform equally well in capturing the posterior variance given by the PF and their performance is much better than the EKF.

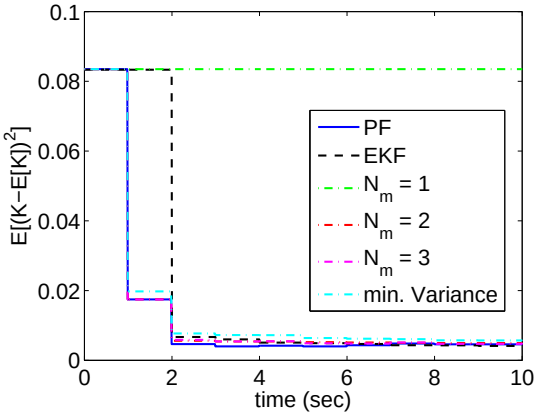
Furthermore, Fig. 4.1(e) and Fig. 4.1(f) show the performance of applied methods in capturing the third posterior central moment for parameter K and state x, respectively. It is clear that the gPC-Bayes method is not able to capture the third central moment for $N_m < 3$. However, there is a significant improvement in capturing the posterior third central moment assuming the PF approximated third central moment to be reference truth when $N_m \geq 3$. This is due to the fact that for capturing the posterior third central moment, the minimum order of matching moment constraints should be at least three. As expected, both the gPC based minimum variance estimator and the EKF do not perform well in capturing



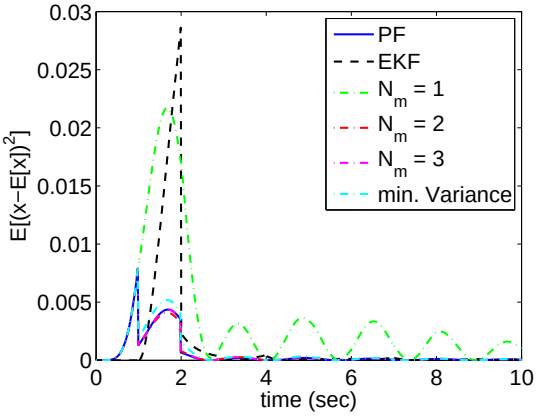
(a) Mean of K



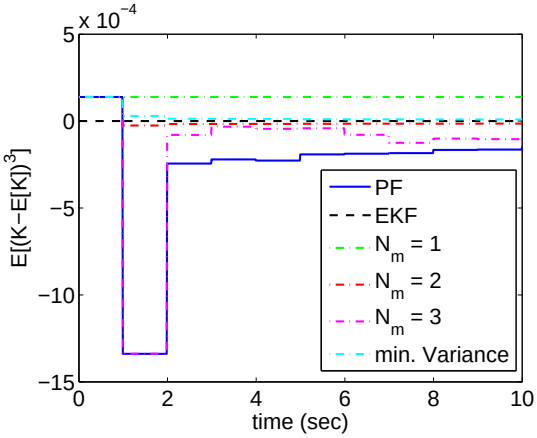
(b) Mean of x



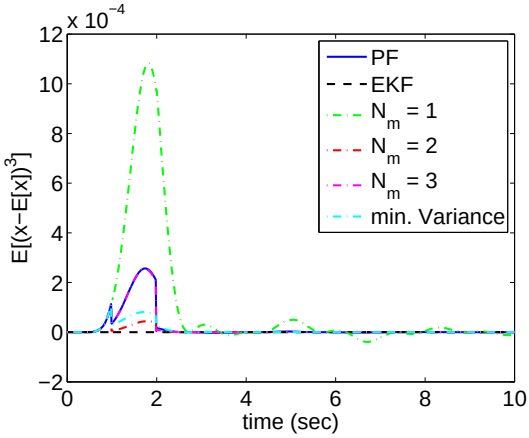
(c) Variance for K



(d) Variance for x



(e) 3^{rd} Central Moment for K



(f) 3^{rd} Central Moment for x

Figure 4.1: Posterior Central Moments for Parameter K and State x for Example 1.

Table 4.1: RMSE error in first three posterior central moments for parameter K.

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	4.1827e+000	2.3343e+000	1.7592e-002
2	3.2244e-001	2.5930e-002	1.4137e-002
3	1.5010e-001	2.5059e-002	3.7661e-003
min. Variance	3.6602e-001	6.9057e-002	1.4929e-002
EKF	4.5070e+000	5.6481e-001	1.4602e-002

Table 4.2: RMSE error in first three posterior central moments for parameter x.

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	5.7567e-001	1.6527e-001	6.9901e-003
2	5.1621e-002	2.2547e-003	1.5859e-003
3	2.0171e-002	8.8121e-004	3.9529e-005
min. Variance	1.3997e-001	8.5417e-003	1.2629e-003
EKF	1.4636e+000	3.6966e-002	1.9292e-003

the third central moment for both K and x.

Finally, Tables 4.1 and 4.2 show the root mean square error over time in capturing central moments for parameter, K and state, x, respectively. We assume the PF estimated posterior central moment to be the reference truth to compute the root mean square error. Although one should be careful about this comparison as the PF does not provide the "truth" posterior moments due to various assumptions involved regarding the selection of importance function in the measurement update part. As expected, the gPC-Bayes method results in less error in estimation of posterior moments for both parameter K and state x as one increases the number of matching moment constraints, i.e., N_m. Also, the gPC based minimum variance estimator performs almost 10 times better than the EKF in estimation of the first two central moments for x.

In summary, it is clear that the proposed methods perform very well as compared to the PF results in capturing not only the posterior mean but also the higher moments. The main advantage of the gPC-Bayes approach is that one can vary the number of moment matching constraints depending upon the desired accuracy in capturing higher order posterior mo-

ments. The poor performance of the EKF algorithm can be attributed to the nonlinearity involved due to simultaneous state and parameter estimation problem.

4.2 Duffing Oscillator

We next consider the Duffing oscillator of Eq. (2.44) in section 2.3.2

$$\ddot{x} + \eta\dot{x} + \alpha x + \beta x^3 = \sin(3t) \tag{4.4}$$

We consider two different scenarios: 1) we consider pure state estimation problem by assuming initial conditions to be uncertain and 2) We consider simultaneous sate and parameter estimation problem by assuming parameters (α and β) to be uncertain.

4.2.1 Second Example: Pure State Estimation

For simulation purposes, nominal parameter values are assumed to be given as:

$$\eta = 1.3663,\ \alpha = -1.3761\ \beta = 2$$

The initial states are assumed to be normally distributed:

$$x(0) = \mathcal{N}(x_0| - 1, 0.25), \quad \dot{x}(0) = \mathcal{N}(\dot{x}_0| - 1, 0.25)$$

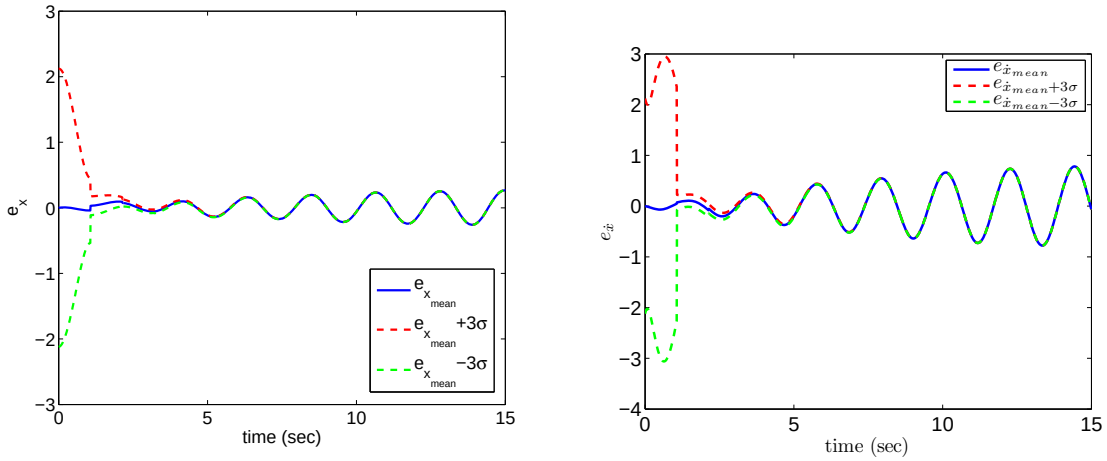
To analyze the effect of initial condition uncertainty, 4^{th} order gPC expansion is considered. Hence, polynomial chaos expansion of states will be equal to:

$$x(0, \boldsymbol{\xi}) = \sum_{k=0}^{4} x_k(0)\psi_k(\boldsymbol{\xi}) \quad x_0(0) = -1, \quad x_1(0) = 0.5, \quad x_k(0) = 0, \quad \text{for} \quad k > 2 \tag{4.5}$$

$$\dot{x}(0, \boldsymbol{\xi}) = \sum_{k=0}^{4} \dot{x}_k(0)\psi_k(\boldsymbol{\xi}) \quad \dot{x}_0(0) = -1, \quad \dot{x}_1(0) = 0.5, \quad \dot{x}_k(0) = 0 \quad \text{for} \quad k > 2 \tag{4.6}$$

where, $\boldsymbol{\xi} = [\xi_1\ \xi_2]^T$ is a vector of normally distributed random variables ξ_1 and ξ_2. Also, $\psi_k(\boldsymbol{\xi})'s$ are Hermite polynomials which are used to describe Gaussian distribution of states.

To verify efficiency of our method, we compared the performance of the proposed methods with the EKF results. The measurement data is assumed to be available at a sampling frequency of 1Hz. A random sample of initial conditions is taken from prior initial condition



(a) $x(t)$ (b) $\dot{x}(t)$

Figure 4.2: Estimation Error and 3σ Bounds for the EKF Approximated Posterior Mean for Example 2

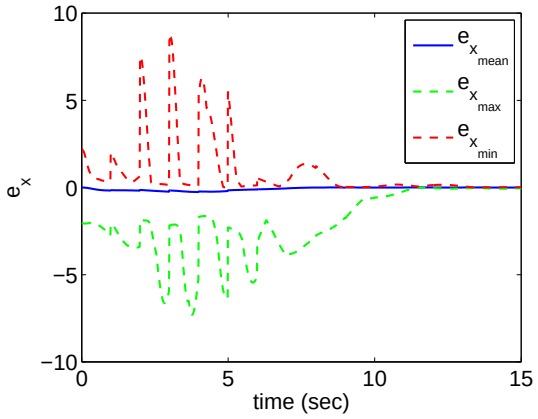
distribution to generate the true measurement data. The true measurement data is then corrupted with a Gaussian white noise of zero mean and variance being:

$$R = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$$

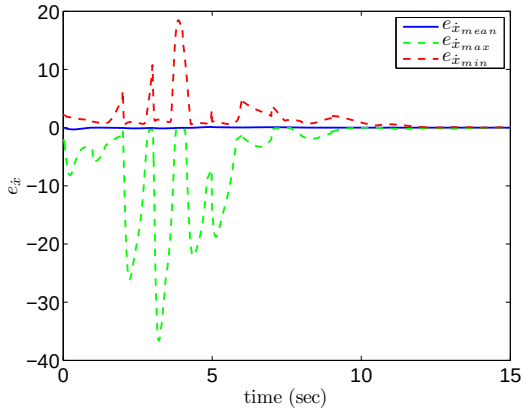
σ is assumed to be 0.05 in our simulations.

Fig. 4.2(a) and Fig. 4.2(b) show the state estimation error for x and $\dot{x}$, respectively. The solid blue line represents the difference between the true value and its mean estimate. Dashed green line shows -3σ bound while the dashed red line represents the 3σ bound. From these plots, it is clear that the state estimation error increases significantly during the time although it is always bounded by $\pm 3\sigma$ bounds. The poor performance the EKF can be attributed to strong nonlinearities and sparse data frequency of 1 Hz.

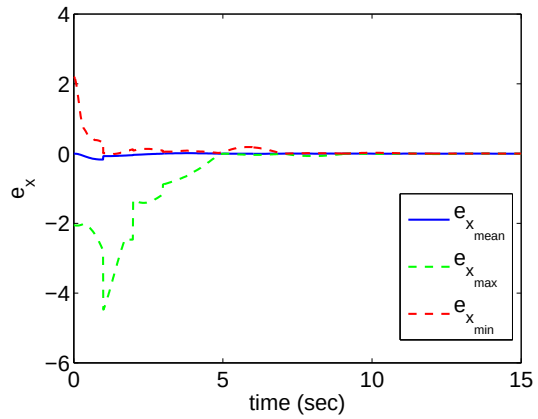
Fig. 4.3 shows the error in state estimates using the gPC-Bayes method for various values of N_m. The solid blue line represents the difference between the true value and its mean estimate. Dashed green line and dashed red line represent the min and max bounds on estimation errors, respectively. It is clear that estimation error and corresponding min-max bounds for estimation error converge to zero over the time. This is due to the fact that posterior density function finally converges to a dirac-delta function around the truth which is expected as number of measurements increases over the time. Also, it should be



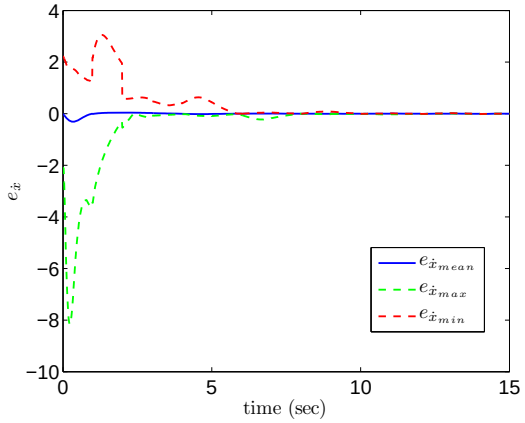
(a) Estimation Error for x ($N_m = 1$)



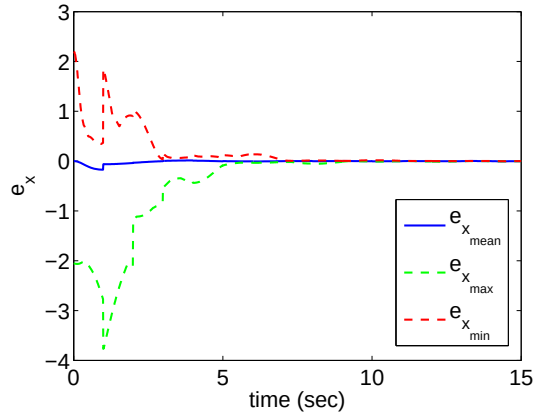
(b) Estimation Error for $\dot{x}$ ($N_m = 1$)



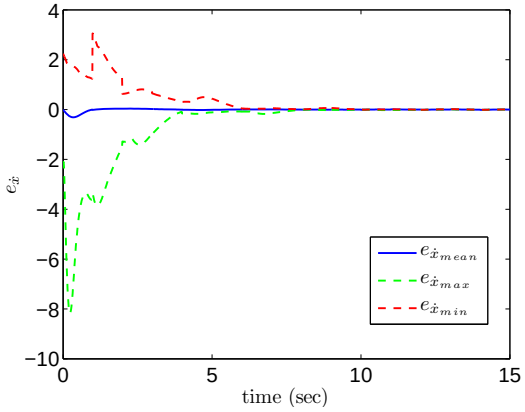
(c) Estimation Error for x ($N_m = 2$)



(d) Estimation Error for $\dot{x}$ ($N_m = 2$)

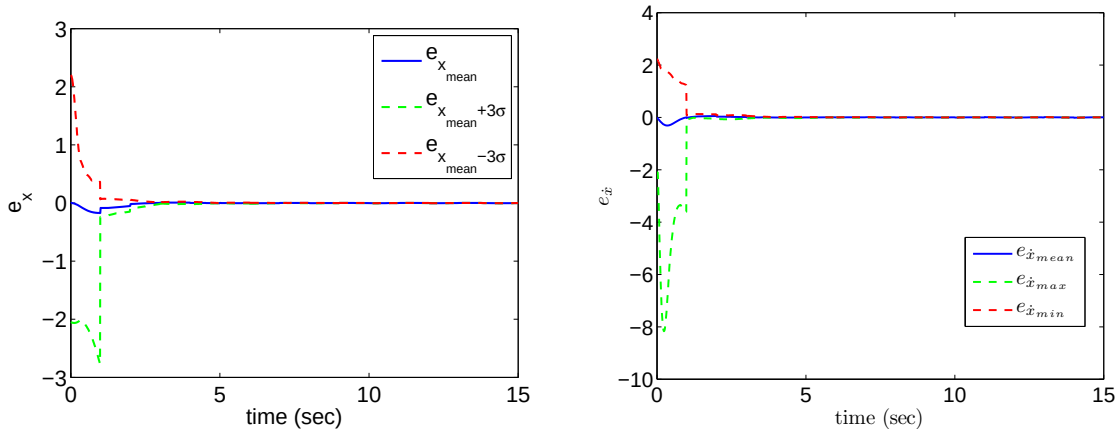


(e) Estimation Error for x ($N_m = 3$)



(f) Estimation Error for $\dot{x}$ ($N_m = 3$)

Figure 4.3: Estimation Error and min-max Bounds for the gPC-Bayes Approximated Posterior Mean for Example 2



(a) $x(t)$ (b) $\dot{x}(t)$

Figure 4.4: Estimation Error and $3-\sigma$ Bounds for the gPC Based Minimum Variance Estimator Approximated Posterior Mean for Example 2

noticed that min-max bounds becomes more and more tighter as one increases the number of matching moment constraints, i.e., N_m. Futhermore, Fig. 4.4 shows the error in state estimates along with its min-max bounds using the gPC based minimum variance estimator. Once again, the estimation error along with min-max bounds converge to zero over the time which can be again attributed to the posterior density function being a delta function as number of measurements increases. Form these results, it is clear that the proposed methods do very well in not only estimating the posterior mean but posterior density function also.

4.2.2 Example 3: Simultaneous State and Parameter Estimation

As the third example, let us consider the problem of simultaneous state and parameter estimation problem for the Duffing oscillator:

$$\ddot{x} + \eta\dot{x} + \alpha x + \beta x^3 = U_{in} \quad x(0) = -1, \quad \dot{x}(0) = -1,\ U_{in} = sin(3t) \tag{4.7}$$

Like in Chapter 2, η and α are assumed to be uniformly distributed uncertain parameters over the intervals $[0.9, 1.4]$ and $[-1.45, -0.95]$, respectively. The nominal value for β is assumed to be 2. For simulation purposes, measurement data is assumed to be available at a sampling frequency of 1Hz. A random sample of η and α is taken from their prior distributions to generate the true measurement data. The true measurement data is corrupted with

a Gaussian white noise of zero mean and variance being 0.05. To represent uncertainty in state and parameter, 7^{th} order gPC expansion is considered and total simulation time interval is assumed to be 5 sec. The initial gPC expansion for η, α and states can be written as:

$$x(0,\boldsymbol{\xi}) = \sum_{k=0}^{7} x_k(0)\phi_k(\boldsymbol{\xi}) \quad x_0(0) = -1 \quad x_k(0) = 0 \quad for \quad k \geq 2 \tag{4.8}$$

$$\dot{x}(0,\boldsymbol{\xi}) = \sum_{k=0}^{7} \dot{x}_k(0)\phi_k(\boldsymbol{\xi}) \quad \dot{x}_0(0) = -1 \quad \dot{x}_k(0) = 0 \quad \text{for} \quad k \geq 2 \tag{4.9}$$

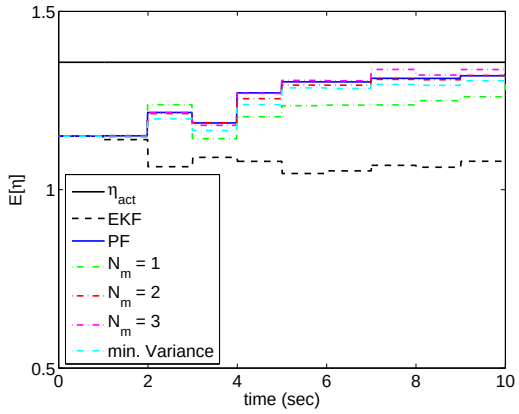
$$\eta(\boldsymbol{\xi}) = \sum_{k=0}^{7} \eta_k\phi_k(\boldsymbol{\xi}) \quad \eta_0 = 1.15, \quad \eta_1 = 0.25 \quad \text{and} \quad \eta_k = 0 \quad \text{for} \quad k > 2 \tag{4.10}$$

$$\alpha(\boldsymbol{\xi}) = \sum_{k=0}^{7} \alpha_k\phi_k(\boldsymbol{\xi}) \quad \alpha_0 = -1.2, \quad \alpha_1 = 0.25 \quad \text{and} \quad \alpha_k = 0 \quad \text{for} \quad k > 2 \tag{4.11}$$

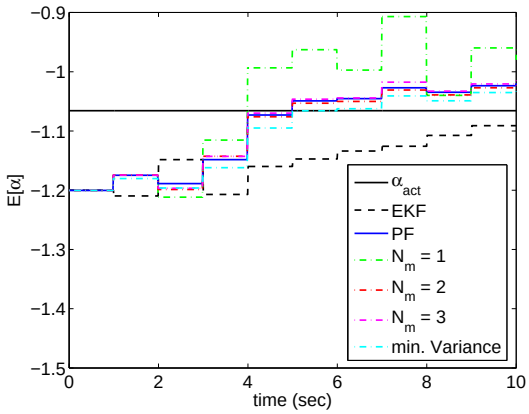
All expectation integrals involved in the gPC-Bayes method are evaluated by using 25 quadrature points in each dimension of random variables ξ_1 and ξ_2 resulting in a total of 625 quadrature points.

Fig. 4.5(a) and Fig. 4.5(b) show posterior mean estimates corresponding to different filters for parameter η and α, respectively. As expected the gPC-Bayes method estimates converges to the PF approximated posterior mean as we increase number of matching moments constraints, i.e., N_m. Also, the gPC based minimum variance estimator performs similar to the gPC-Bayes method with $N_m = 2$. It is also clear that the EKF performs worst in approximating the posterior mean for parameters.

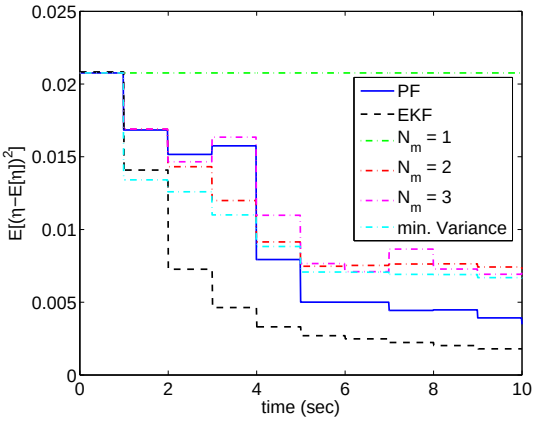
The posterior variance for parameters η and α has been shown in Fig. 4.5(c) and Fig. 4.5(d), respectively. As expected the gPC-Bayes method does not capture the posterior variance for $N_m = 1$, however, the performance of the gPC-Bayes approach improves significantly for N_m greater than one. Furthermore, Fig. 4.5(e) and Fig. 4.5(f) show the plot of posterior third central moment corresponding to various filters for parameters η and α, respectively. As expected, the gPC based minimum variance estimator and EKF are not able to capture the third central moments assuming the PF approximation to be the reference truth. The gPC-Bayes method performs well in approximating the third central moment up to two seconds for $N_m = 3$. However, the gPC-Bayes method performs poorly even for $N_m = 3$ for time greater than two seconds. The poor performance the gPC-Bayes method can be attributed to the finite gPC approximation. We will discuss this issue in much more detail in section



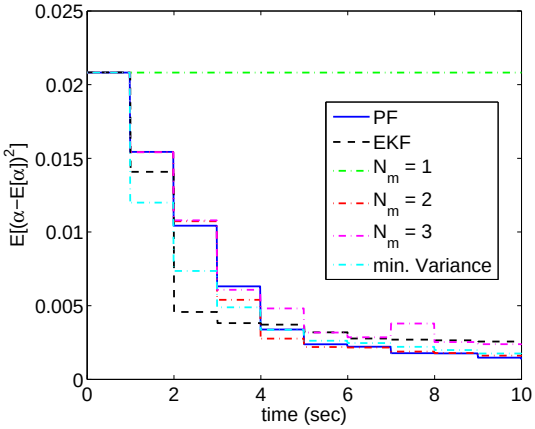
(a) Mean for η



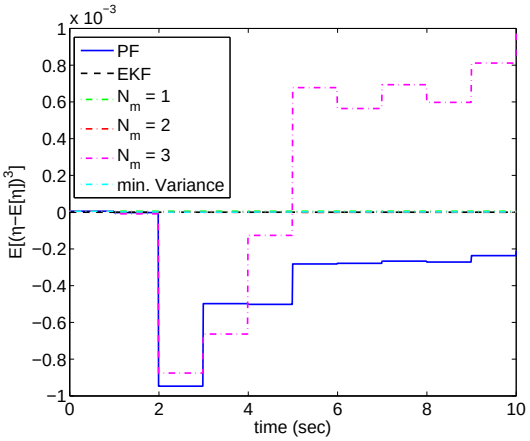
(b) Mean for α



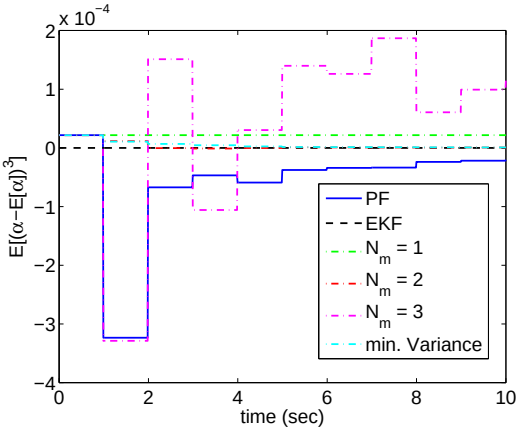
(c) Variance for η



(d) Variance for α



(e) 3^{rd} Central Moment for η



(f) 3^{rd} Central Moment for α

Figure 4.5: Posterior Central Moments for η and α for Example 3.

4.2.2.

Tables 4.3 and 4.4 shows the root mean square error over time in approximating first three posterior central moments for parameters, η and α, respectively. We assume the PF approximated posterior moments to be reference truth for the evaluation of root mean square error. As expected, the gPC-Bayes method results in less error in estimation of posterior moments as one increases the number of matching moment constraints, i.e., N_m. Also, proposed methods performs a order of magnitude better than the EKF in estimation of the first three central moments.

Table 4.3: RMSE error in first three posterior central moments for η

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	1.6894e+000	3.9390e-001	1.3371e-002
2	2.1844e-001	7.8388e-002	1.3233e-002
3	3.2892e-001	7.4929e-002	2.1433e-002
min. Variance	5.8136e-001	8.3721e-002	1.3263e-002
EKF	5.4728e+000	2.6102e-001	1.3245e-002

Table 4.4: RMSE error in first three posterior central moments for α

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	1.9049e+000	4.9394e-001	3.8996e-003
2	1.4937e-001	1.1782e-002	3.5675e-003
3	1.4523e-001	2.9626e-002	4.3245e-003
min. Variance	4.3090e-001	4.8743e-002	3.5880e-003
EKF	2.5449e+000	1.8308e-001	3.4666e-003

Fig. 4.6 shows the plots for first three posterior central moments for both states x and $\dot{x}$. Also, Tables 4.5 and 4.6 shows the RMSE error over time in approximating first three central moments for x and $\dot{x}$, respectively. Form these results, it is clear that the gPC bayes method performs well in capturing all three central moments as the number of matching moment constraints are increased. The EKF performs worst among all the filters as it even fails to capture the posterior mean also. The gPC based minimum variance filter performs well in capturing the first two central moments but is not able to capture the third central

Table 4.5: RMSE error in first three posterior central moments for x

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	7.0649e-001	8.5432e-002	1.9670e-003
2	6.7726e-002	2.3682e-003	1.5666e-004
3	6.1157e-002	6.2015e-003	3.0797e-004
min. Variance	1.7341e-001	5.7130e-003	1.7355e-004
EKF	9.5327e-001	1.6320e-002	9.2697e-005

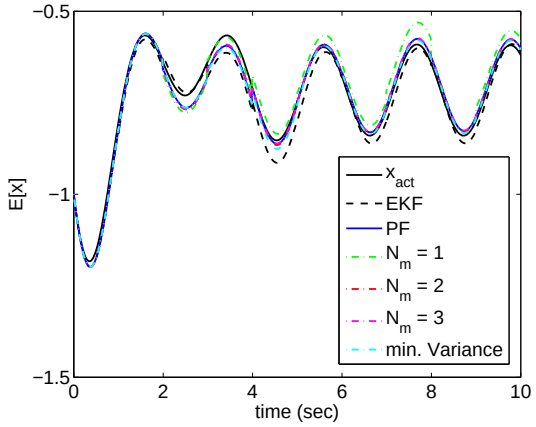
Table 4.6: RMSE error in first three posterior central moments for $\dot{x}$

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	2.8121e-001	1.9142e-002	3.8337e-004
2	3.7524e-002	1.4365e-003	1.0563e-004
3	3.1312e-002	2.0126e-003	7.7402e-005
min. Variance	1.0039e-001	2.6557e-003	1.1128e-004
EKF	7.1761e-001	1.6451e-002	1.7959e-004

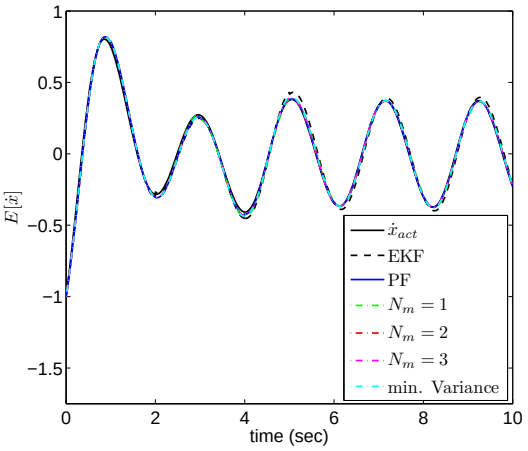
moment accurately. It should be noticed that the error of the gPC based minimum variance estimator is considerably less than the EKF.

Error Analysis

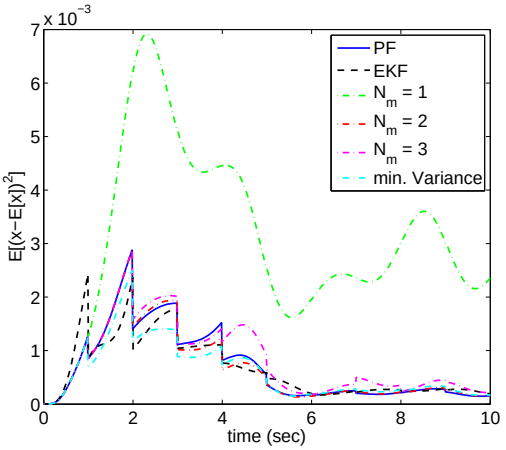
As we noticed in the last section, after some time the posterior distribution/moments approximated by the gPC-Bayes method do not match well with those approximated by the PF especially for the third central moment approximation. In order to analyze this issue in more detail, let us reconsider Fig. 4.6. Notice that all three posterior central moments for x approximated by the gPC-Bayes method match well with the PF approximation up to first measurement update, i.e., $t = 1$ second for $N_m = 3$. To make this point more clear, the first three posterior central moments for x at $t = 1$ second are listed in Table 4.7. Furthermore, Fig. 4.7 shows histograms corresponding to posterior distribution for x after the first measurement update at $t = 1$ second for gPC-Bayes ($N_m = 3$) and the PF. From Fig. 4.7, it is clear that even though the first three central moments of x are the same at $t = 1$ second, but the posterior distributions approximated by using the PF and gPC-Bayes method are



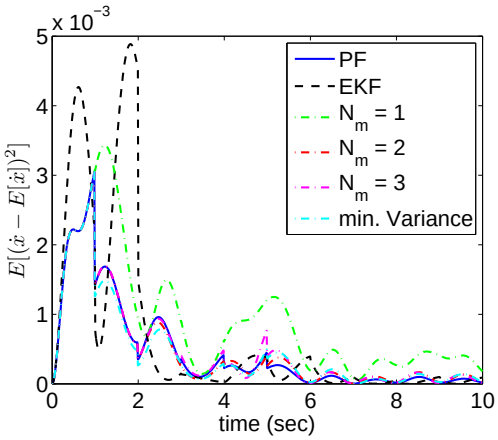
(a) Mean for x



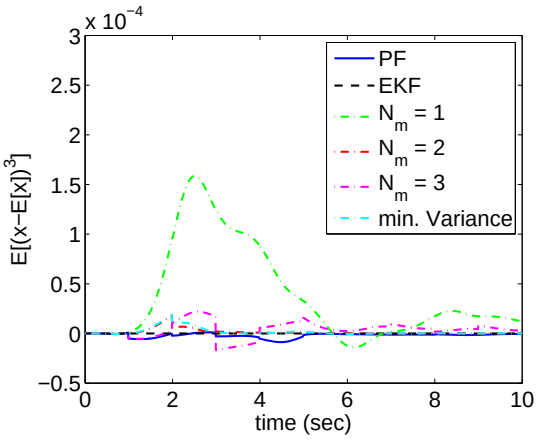
(b) Mean for $\dot{x}$



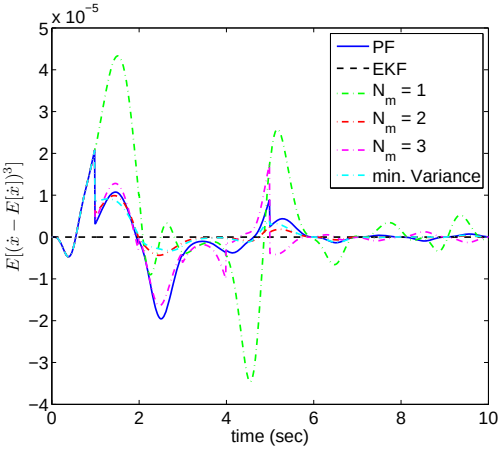
(c) Variance for x



(d) Variance for $\dot{x}$



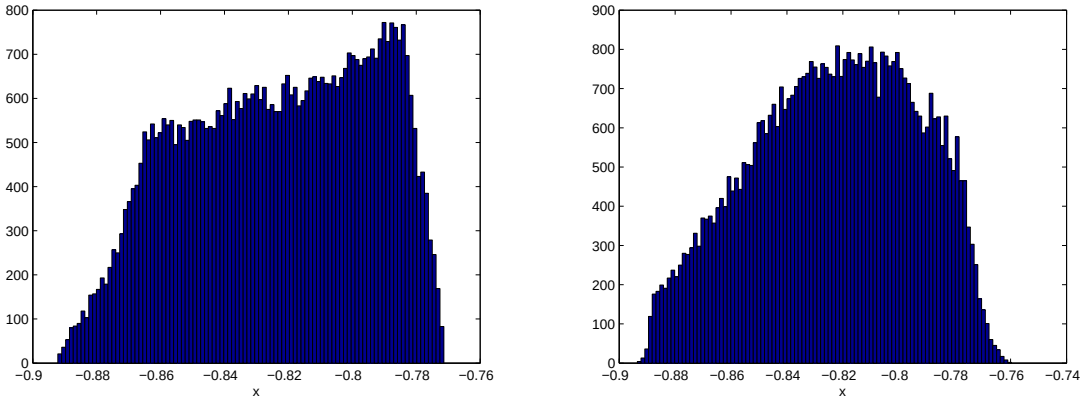
(e) 3^{rd} Central Moment for x



(f) 3^{rd} Central Moment for $\dot{x}$

Figure 4.6: Posterior Central Moments for States for Example 3.

bit different due to mismatch in higher order moments. This discrepancy between the PF and the gPC-Bayes method grows when we propagate these two distributions to the next measurement update interval as shown in histogram plots of Fig. 4.8 and moment data in Table 4.8. Ideally one can overcome this error by increasing the number of moment matching constraints and the order of the gPC expansion which leads to higher computational load. In practice, one needs to compromise between the computational load and accuracy in approximating higher order moments. Currently, the computational load measured in terms of CPU time in running the MATLAB based simulation is minimal as shown in Table 4.9.



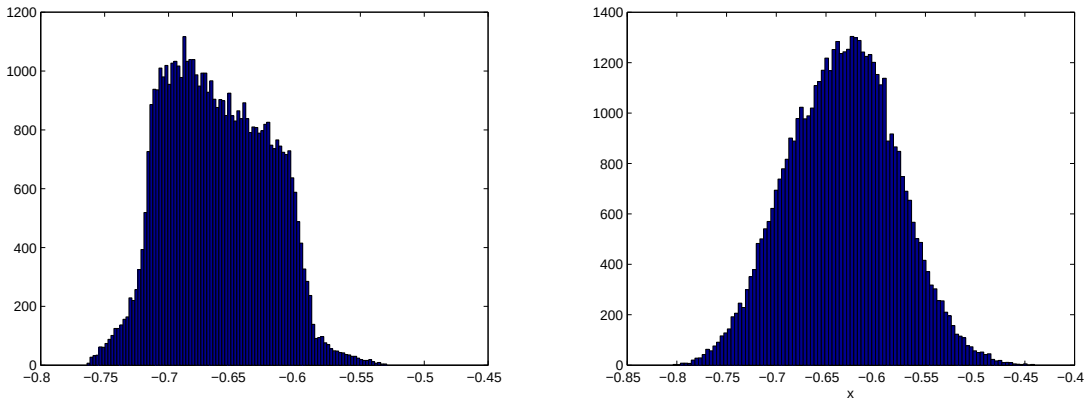
(a) gPC-Bayes ($N_m = 3$)

(b) Particle Filter

Figure 4.7: histograms of state x_1 after the measurement update at $t = 1$ *sec.*

Table 4.7: First three posterior central moments of state x at $t = 1$ *sec.* by using the PF and gPC-Bayes method ($N_m = 3$)

Method	Mean	2^{nd} Central Moment	3^{rd} Central Moment
PF	-8.2328e-001	8.5817e-004	-5.0996e-006
gPC-Bayes ($N_m = 3$)	-8.2322e-001	8.5724e-004	-5.0696e-006



(a) gPC-Bayes ($N_m = 3$)

(b) Particle Filter

Figure 4.8: histograms of state x before the measurement update at $t = 2$ *sec.*

Table 4.8: First three central moments of state x before the measurement update at $t = 2$ *sec.* by using PF and gPC-Bayes method ($N_m = 3$)

Method	Mean	2^{nd} Central Moment	3^{rd} Central Moment
PF	-6.3336e-001	2.9261e-003	1.7150e-006
gPC-Bayes ($N_m = 3$)	-6.6187e-001	1.5869e-003	1.0685e-006

Table 4.9: Computational time required for different estimation approaches for Example 3

N_m	EKF	PF	min. Variance	gPC-Bayes
1	0.955887	3.6260e+004	2.1469e+001	2.0219e+002
2				2.8672e+002
3				1.6657e+004

4.3 Example 4: Hovering Helicopter Model

As the last example, we examine efficiency of proposed approach on hovering helicopter model considered in section 2.3.3:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & -g & 0 \\ 1.26 & -1.765 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} - \begin{pmatrix} 0.086 \\ -7.408 \\ 0 \\ 0 \end{pmatrix} K_{lqr} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \tag{4.12}$$

where, K_{lqr} and initial conditions are given as:

$$K_{lqr} = [1.989 \quad -0.256 \quad -0.7589 \quad 1], \qquad X_{in} = [0.7929 \quad -0.0466 \quad -0.1871 \quad 0.5780]^T$$

Like in Chapter 2, p_1 and p_2 are assumed to be uniformly distributed parameters over the intervals $[-0.2, 0]$ and $[0, 0.2]$, respectively. For simulation purposes, measurement data is assumed to be available at a sampling frequency of 1Hz. A random sample of unknown parameters is taken from their prior distributions to generate the true measurement data. The true measurement data is corrupted with a Gaussian white noise of zero mean and standard deviation being 0.15 times an identity matrix. To represent uncertainty in state and parameter, 7^{th} order gPC expansion is considered and total simulation time interval is assumed to be 10 sec. The initial gPC expansion for parameters and states can be written as:

$$x_i(0, \boldsymbol{\xi}) = \sum_{k=0}^{7} x_{i_k}(0)\phi_k(\boldsymbol{\xi}) \quad x_{i_k}(0) = 0 \quad \text{for} \quad k \geq 2 \tag{4.13}$$

$$p_1(\boldsymbol{\xi}) = \sum_{k=0}^{7} p_{1_k}\phi_k(\boldsymbol{\xi}) \quad p_{1_0} = -0.1, \quad p_{1_1} = 0.1 \quad \text{and} \quad p_{1_k} = 0 \quad \text{for} \quad k > 2 \tag{4.14}$$

$$p_2(\boldsymbol{\xi}) = \sum_{k=0}^{7} p_{2_k}\phi_k(\boldsymbol{\xi}) \quad p_{2_0} = 0.1, \quad p_{2_1} = 0.1 \quad \text{and} \quad p_{2_k} = 0 \quad \text{for} \quad k > 2 \tag{4.15}$$

where,

$$x_{1_0}(0) = 0.7929, \quad x_{2_0}(0) = -0.0466, \quad x_{3_0}(0) = -0.1871, \quad x_{4_0}(0) = 0.5780$$

Fig. 4.9, Fig. 4.10 and Fig. 4.11 show first three posterior moments corresponding to various filters for parameters and states, respectively. Furthermore, the RMSE error over

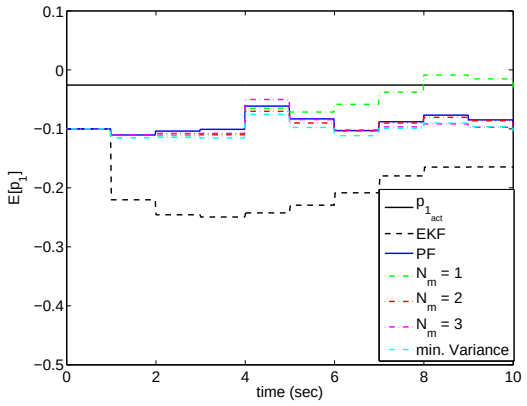
Table 4.10: RMSE error in first three posterior central moments for p_1

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	1.1953e+000	2.9702e-002	1.1897e-003
2	1.5290e-001	1.3637e-002	1.2369e-003
3	2.5390e-001	2.5835e-002	2.9821e-003
min. Variance	3.5699e-001	1.2942e-002	1.2012e-003
EKF	3.7481e+001	8.3300e-002	1.1908e-003

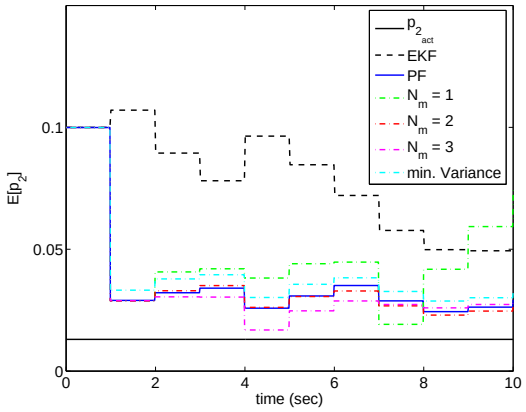
Table 4.11: RMSE error in first three posterior central moments for p_2

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	4.5432e-001	9.2948e-002	3.3328e-005
2	3.9870e-002	1.0262e-003	3.8086e-005
3	1.3396e-001	2.5424e-003	7.3500e-005
min. Variance	1.3444e-001	1.6340e-003	4.3682e-005
EKF	1.8933e+001	4.0457e-002	4.9844e-005

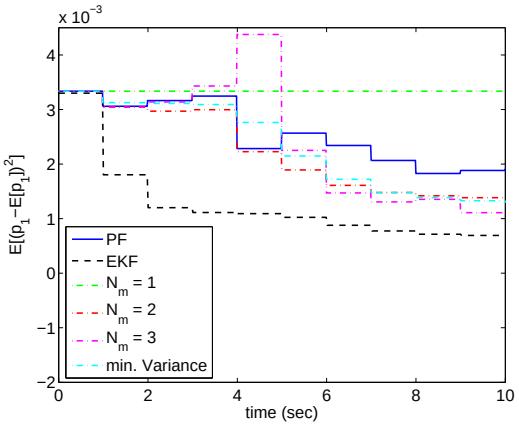
time in approximating first three central moments for parameters and states are listed in Tables 4.10-4.15. As expected, the accuracy of the gPC bayes method in capturing first three central moments improves as the number of matching moment constraints are increased. Like in the previous example, the performance of the gPC-Bayes filter in approximating the third central moment degrades over the time which can be attributed to finite order gPC approximation. The EKF performs worst among all the filters as it even fails to capture the posterior mean. The gPC based minimum variance filter performs better than the EKF in capturing the first two central moments.



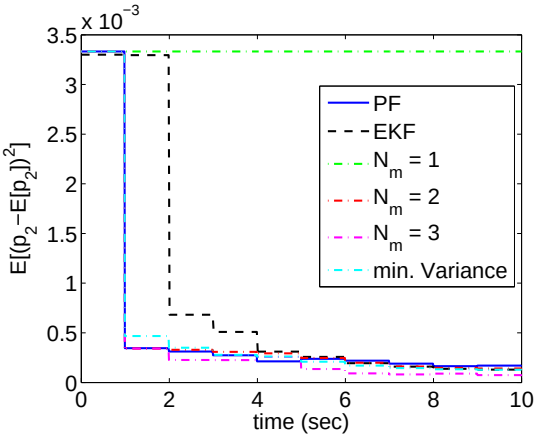
(a) Mean of p_1



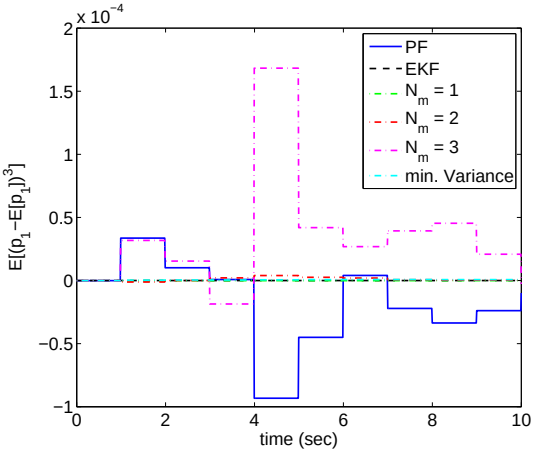
(b) Mean of p_2



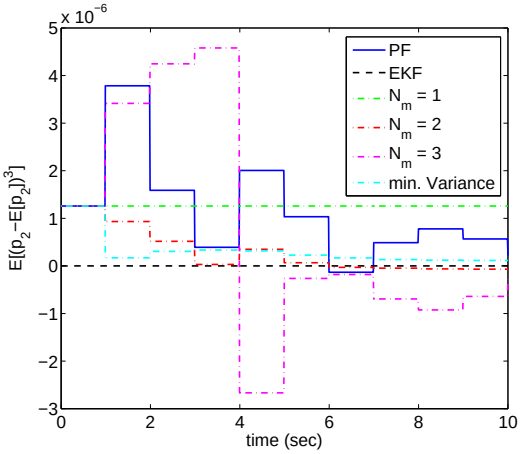
(c) Variance of p_1



(d) Variance of p_2

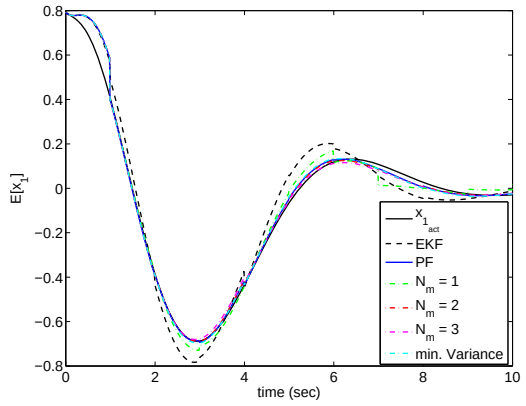


(e) 3^{rd} Central Moment of p_1

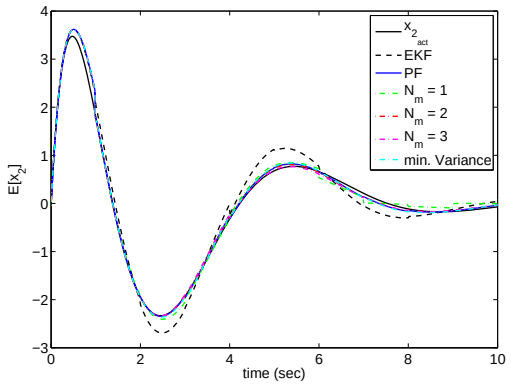


(f) 3^{rd} Central Moment of p_2

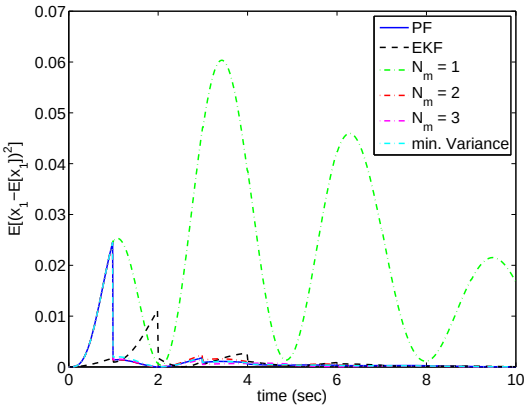
Figure 4.9: Posterior Central Moments for Parameters for Example 4.



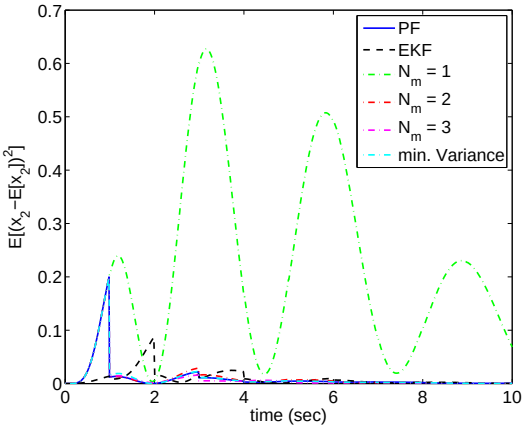
(a) Mean of x_1



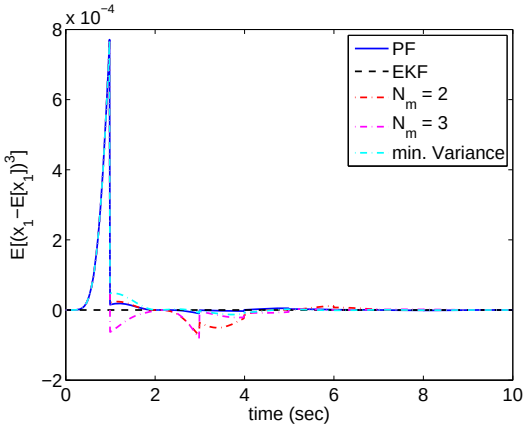
(b) Mean of x_2



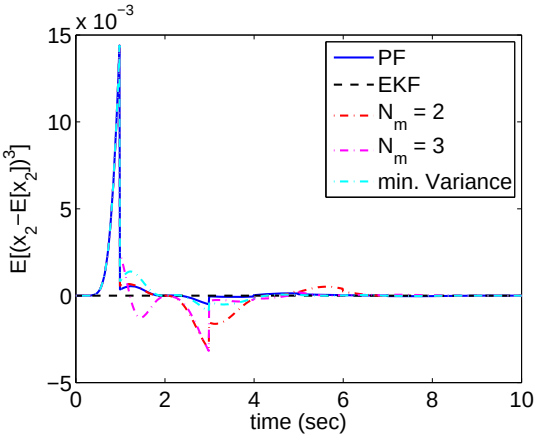
(c) Variance of x_1



(d) Variance of x_2

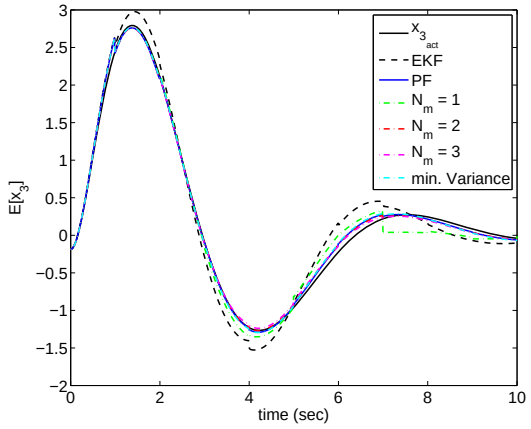


(e) 3^{rd} Central Moment of x_1

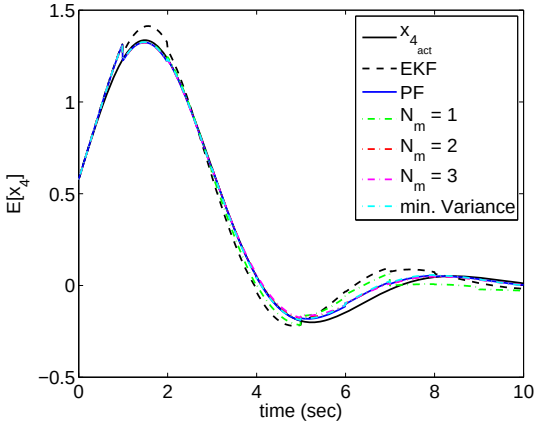


(f) 3^{rd} Central Moment of x_2

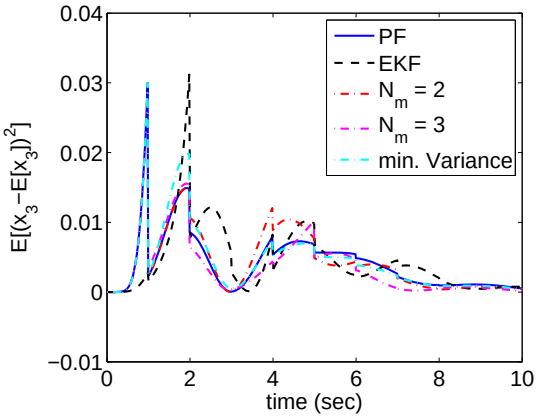
Figure 4.10: Posterior Central Moments for States (x_1 and x_2) for Example 4.



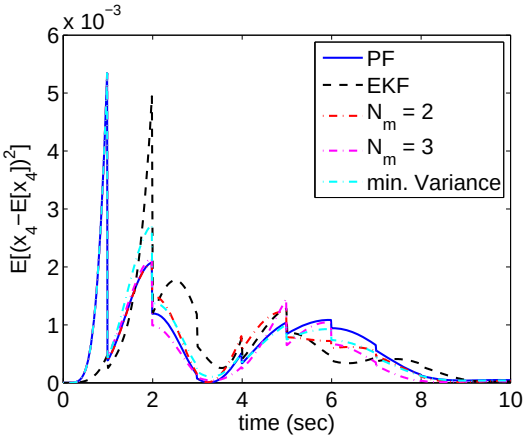
(a) Mean of x_3



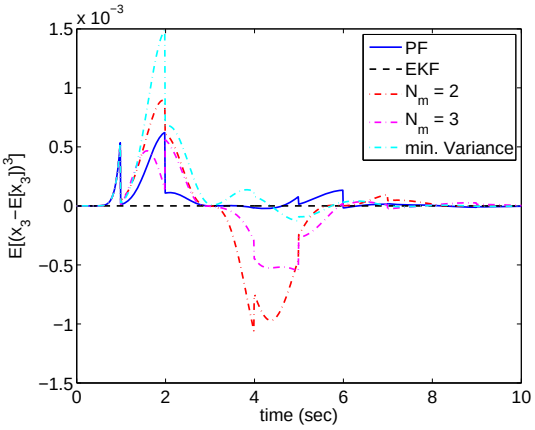
(b) Mean of x_4



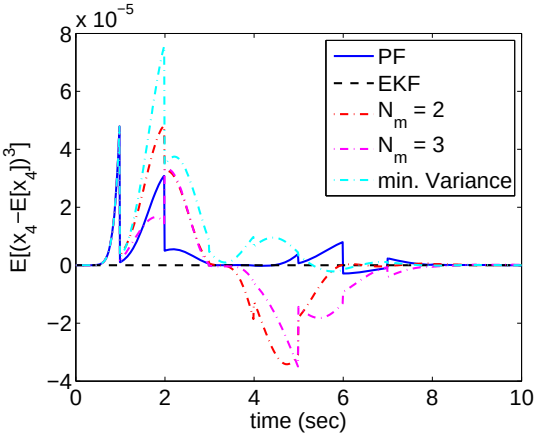
(c) Variance of x_3



(d) Variance of x_4



(e) 3^{rd} Central Moment of x_3



(f) 3^{rd} Central Moment of x_4

Figure 4.11: Posterior Central Moments for States (x_3 and x_4) for Example 4.

Table 4.12: RMSE error in first three posterior central moments for x_1

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
N_m	e_{m_1}	e_{m_2}	e_{m_3}
1	7.3690e-001	8.2452e-001	1.6527e-001
2	9.0236e-002	5.6440e-003	5.0475e-004
3	2.6761e-001	6.2849e-003	5.5725e-004
min. Variance	1.0965e-001	4.0795e-003	2.1633e-004
EKF	4.9379e-001	1.0192e-001	2.5809e-003

Table 4.13: RMSE error in first three posterior central moments for x_2

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	2.1177e+000	8.4853e+000	4.8704e+000
2	2.6311e-001	6.0901e-002	1.7531e-002
3	8.8713e-001	7.1439e-002	1.6858e-002
min. Variance	3.4365e-001	4.0294e-002	6.6701e-003
EKF	1.9873e+000	7.6455e-001	4.3517e-002

Table 4.14: RMSE error in first three posterior central moments for x_3

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	3.0801e+000	7.0409e+000	4.1151e+000
2	3.5989e-001	4.2817e-002	1.0236e-002
3	7.3013e-001	3.8105e-002	6.5463e-003
min. Variance	2.9859e-001	3.4430e-002	6.4499e-003
EKF	1.7666e+000	1.2186e-001	3.9159e-003

Table 4.15: RMSE error in first three posterior central moments for x_4

N_m	Mean	2^{nd} Central Moment	3^{rd} Central Moment
1	1.0644e+000	7.8987e-001	1.5967e-001
2	1.3487e-001	5.3220e-003	3.8942e-004
3	2.4717e-001	4.5880e-003	3.5806e-004
min. Variance	1.0671e-001	5.1191e-003	3.7074e-004
EKF	8.0420e-001	2.1144e-002	2.2017e-004

Finally, Table 4.16 shows the CPU time required for Particle Filter, EKF, gPC-Bayes, and gPC based minimum variance estimator implementation in the MATLAB environment. Offcourse, the EKF is faster than all the other approaches, but similar to other examples, because of low measurement data frequency and also nonlinearities of augmented system, the EKF also results in poor estimates for both states and parameter. The gPC-Bayes method with low values of N_m is much faster than the PF. However by increasing the number of matching moment constraints (N_m), computational time increases. Also, the gPC based minimum variance estimator performs much faster than the gPC-Bayes method but does not capture higher order moments as well as the gPC-Bayes method.

Table 4.16: Computational time (seconds) required for different estimation approaches

N_m	EKF	PF	min. Variance	gPC-Bayes
1	1.1454	8.8284e+003	2.5619e+001	2.6441e+002
2				6.0449e+002
3				3.4292e+004

Chapter 5

Conclusions

In this thesis, two new recursive approaches have been developed to provide accurate estimates for posterior moments of both parameters and system states while making use of the generalized polynomial chaos framework for uncertainty propagation. The main idea of the Generalized polynomial chaos method is to expand random state and input parameter variables involved in a stochastic differential/difference equation in a polynomial expansion. These polynomials are associated with the prior pdf for the input parameters. Later, Galerkin projection is used to obtain deterministic system of equations for the expansion coefficients. The first proposed approach (gPC-Bayes) provides means to update prior expansion coefficients by constraining the polynomial chaos expansion to satisfy the desired number of posterior moment constraints derived from the Bayes' rule. The second proposed approach makes use of the minimum variance formulation to update polynomial chaos expansion coefficients. The main advantage of proposed methods is that they not only provide point estimate for the state and parameters but they also provide statistical confidence bounds associated with these estimates.

Four different numerical examples ranging from one dimensional system to four dimensional system are considered to compare the performance of proposed methods against benchmark algorithms like the extended Kalman filter and the particle filter. The numerical results show that the proposed methodologies perform much better than the extended Kalman filter in capturing the posterior mean for both state and parameter. Furthermore, it is demonstrated that one can converges to the particle filter estimates for not only posterior mean but also higher order moments by increasing the number of matching moment constraints in the

gPC-Bayes method. The minimum variance method is shown to provide consistently a valid estimates for both posterior mean and variance for both state and parameters. Although we have not performed a detailed computational analysis, numerical example results show that the processor time associated with the polynomial chaos based minimum variance estimator and the gPC-Bayes method with one or two matching moment constraint is much lower than that associated with the particle filter and slightly higher than the extended Kalman filter while consistently providing accurate estimates for posterior mean and variance. Like any other nonlinear filtering approach, the computational burden increases considerably as one increases the number of matching moment constraints which helps in providing a better spectral content of the posterior density function.

An open research issue is to associate the error in approximating moments with the order of the polynomial chaos expansion. This kind of error analysis can help one in selecting the order of polynomial chaos expansion to match desired order of moments. However, this analysis is difficult due to the absence of any closure in moment space. Furthermore, another important issue is the consideration of the white noise in the polynomial chaos framework. Although some recent progress has been made in this direction but a lot more needs to be done. It is hoped that further research in this area will shed light on these problems.

Bibliography

[1] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems 1," *Transactions of the ASME–Journal of Basic Engineering*, Vol. 82, No. Series D, 1960, pp. 35–45.

[2] Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Dover, New York, 1970.

[3] Gelb, A., *Applied Optimal Estimation*, MIT Press, 1974, 203-214.

[4] Schmidt, S. F., "Application of State-Space Methods to Navigation Problems," *Advanced Control Systems*, Vol. 3, 1966, pp. 293–340.

[5] Sayed, A., "A framework for state-space estimation with uncertain models," *IEEE Transactions on Automatic Control*, Vol. 46, No. 7, Jul 2001, pp. 998–1013.

[6] Nagpal, K. M. and Khargonekar, P. P., "Filtering and Smoothing in an $\mathcal{H}_\infty$ Setting," *IEEE Transactions on Automatic Control*, Vol. AC-36, No. 2, 1991, pp. 152–166.

[7] Simon, D., *Optimal state estimation: Kalman, H_∞ and nonlinear approaches*, John Wiley and Sons, 2006.

[8] Ian R. Petersen, A. V. S., *Robust Kalman filtering for signals and systems with large uncertainties*, Springer, 1999.

[9] Polyak, B. T., Nazin, S. A., Durieu, C., and Walter, E., "Ellipsoidal parameter or state estimation under model uncertainty," *Automatica*, Vol. 40, No. 7, 2004, pp. 1171 – 1179.

[10] Xie, L., Soh, Y. C., and de Souza, C., "Robust Kalman filtering for uncertain discrete-time systems," *IEEE Transactions on Automatic Control*, Vol. 39, No. 6, 1994, pp. 1310–1314.

[11] Anderson, B. D. O. and Moore, J. B., *Optimal Filtering*, Prentice-Hall, 1979, 277-297.

[12] Julier, S. J., "New extension of the Kalman filter to nonlinear systems," *Proceedings of SPIE*, 1997, pp. 182–193.

[13] Daum, F. E., "Bounds on performance for multiple target tracking," *Automatic Control, IEEE Transactions on*, Vol. 35, No. 4, Apr 1990, pp. 443–446.

[14] Daum, F., Yau, S.-T., and Yau, S., "Comments on "Finite-dimensional filters with nonlinear drift" [and addendum]," *Aerospace and Electronic Systems, IEEE Transactions on*, Vol. 34, No. 2, Apr 1998, pp. 689–692.

[15] Daum, F. E., "Exact solution to the Zakai equation for certain diffusions," *Decision and Control, 1985 24th IEEE Conference on*, Vol. 24, Dec. 1985, pp. 1964–1965.

[16] Daum, F. E., "A new nonlinear filtering formula for discrete time measurements," *Decision and Control, 1985 24th IEEE Conference on*, Vol. 24, Dec. 1985, pp. 1957–1958.

[17] Daum, F. E., "A new nonlinear filtering formula non-Gaussian discrete time measurements," *Decision and Control, 1986 25th IEEE Conference on*, Vol. 25, Dec. 1986, pp. 1030–1031.

[18] Daum, F. E., "Exact finite dimensional nonlinear filters," *Decision and Control, 1985 24th IEEE Conference on*, Vol. 24, Dec. 1985, pp. 1938–1945.

[19] Daum, F. E., "Exact finite dimensional nonlinear filters for continuous time processes with discrete time measurements," *Decision and Control, 1984. The 23rd IEEE Conference on*, Vol. 23, Dec. 1984, pp. 16–22.

[20] Ito, K.; Xiong, K., "Gaussian filters for nonlinear filtering problems," *Automatic Control, IEEE Transactions on*, Vol. 45, No. 5, 2000, pp. 910–927.

[21] Julier, S. and Uhlmann, J., "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, Vol. 92, No. 3, 2004, pp. 401–422.

[22] Sloan, I. H. and Woniakowski, H., "When Are Quasi-Monte Carlo Algorithms Efficient for High Dimensional Integrals?" *Journal of Complexity*, 1998, pp. 1–33.

[23] "Special Issue on Monte Carlo Methods for Statstical Signal Processing," *IEEE Trasactions on Signal Processing*, Vol. 50, No. 2, 2002.

[24] Daum, F. and Huang, J., "Curse of dimensionality and particle filters," *Aerospace Conference, 2003. Proceedings. 2003 IEEE*, Vol. 4, March 8-15, 2003, pp. 1979–1993.

[25] Arulampalam, M., Maskell, S., Gordon, N., and Clapp, T., "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, Vol. 50, No. 2, 2002, pp. 174–188.

[26] Nygaard, J., Henrik, N., and Young, P. C., "Parameter Estimation in Stochastic Differential Equations: an Overview," *Annual Reviews in Control*, Vol. 24, 2000, pp. 83–94.

[27] Wiener, N., "The Homogeneous Chaos," *American Journal of Mathematics*, Vol. 60, No. 4, Oct. 1938, pp. 897–936.

[28] Xiu, D. and Karniadakis, G. E., "The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations," *SIAM Journal on Scientific Computing*, Vol. 24, No. 2, 2002, pp. 619–644.

[29] Askey, R. and Wilson, J., "Some Basic Hypergeometric Polynomials that Generalize Jacobi Polynomials," *Memoirs Amer. Math. Soc.*, Vol. 319, 1985.

[30] Ghanem, R. G. and Spanos, P. D., *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, NY, 1991.

[31] Blanchard, E., "Parameter Estimation Method using an Extended Kalman Filter," *Mechanical Engineering*, 2007.

[32] Li, J. and Xiu, D., "A generalized polynomial chaos based ensemble Kalman filter with high accuracy," *Journal of Computational Physics*, Vol. 228, No. 15, Aug. 2009, pp. 5454–5469.

[33] Pence, B. L., Fathy, H. K., and Stein, J. L., "A Maximum Likelihood Approach To Recursive Polynomial Chaos Parameter Estimation," *2010 American Control Conference, Marriott Waterfront*, Baltimore, MD, USA, 2010, pp. 2144–2151.

[34] Marzouk, Y. M., Najm, H. N., and Rahn, L. a., "Stochastic spectral methods for efficient Bayesian solution of inverse problems," *Journal of Computational Physics*, Vol. 224, No. 2, June 2007, pp. 560–586.

[35] Marzouk, Y. and Xiu, D., "A Stochastic Collocation Approach to Bayesian Inference in Inverse Problems," *Communications in Computational Physics*, Vol. 6, No. 4, 2009, pp. 826–847.

[36] Blanchard, E., "A Polynomial-chaos based Bayesian approach for estimating uncertain parametersof mechanical systems," *Proceedings of the ASME 2007 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE 2007*, 2007, pp. 1–9.

[37] Dutta, P. and Bhattacharya, R., "Nonlinear Estimation with Polynomial Chaos and Higher Order Moment Updates," *2010 American Control Conference, Marriott Waterfront*, Baltimore, MD, USA, 2010, pp. 3142–3147.

[38] A. Apte, M. Hairer, A. S. and Voss, J., "Sampling the Posterior: An Approach to Non-Gaussian Data Simulation," *Physica D*, Vol. 230, 2007, pp. 50–64.

[39] G. L. Eyink, J. R. and Alexander, F. J., "A Statistical-Mechanical Approach to Data Assimilation for Nonlinear Dynamics: Evolution Approximations," *J. Stat. Phys (in review)*.

[40] Doucet, A., de Freitas, N., and Gordon, N., *Sequential Monte-Carlo Methods in Practice*, Springer-Verlag, April 2001, 6-14.

[41] Iyengar, R. N. and Dash, P. K., "Study of the random vibration of nonlinear systems by the Gaussian closure technique," *Journal of Applied Mechanics*, Vol. 45, 1978, pp. 393–399.

[42] Roberts, J. B. and Spanos, P. D., *Random Vibration and Statistical Linearization*, Wiley, 1990, 122-176.

[43] Lefebvre, T., Bruyninckx, H., and Schutter, J. D., "Kalman Filters of Non-Linear Systems: A comparison of Performance," *International journal of Control*, Vol. 77, No. 7, 2004, pp. 639–653.

[44] Lefebvre, T., Bruyninckx, H., and Schutter, J. D., "Comment on A New Method for the Nonlinear Transformations of Means and Covariances in Filters and Estimators," *IEEE Transactions on Automatic Control*, Vol. 47, No. 8, 2002.

[45] C. Archambeau, D. Cornford, M. O. J. S.-T., "Gaussian Process Approximations of Stochastic Differential Equations," *Journal of Machine Learning Research Workshop and Conference Proceedings*, Vol. 1, 2007, pp. 1–16.

[46] Kullback, S. and Leibler, R., "On Information and Sufficiency," *Ann. Math. Stat.*, Vol. 22, 1951, pp. 79–86.

[47] Wiener, N., "The Homogeneous Chaos," *American Journal of Mathematics*, Vol. 60, No. 4, 1938, pp. 897–936.

[48] Xiu, D. and Karniadakis, G. E., "The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations," *SIAM Journal on Scientific Computing*, Vol. 24, No. 2, 2002, pp. 619.

[49] Dalbey, K., Patra, A. K., Pitman, E. B., Bursik, M. I., and Sheridan, M. F., "Input uncertainty propagation methods and hazard mapping of geophysical mass flows," Vol. 113, 2008, pp. 1–16.

[50] Cheney, E. W. and Kincaid, D., *Numerical mathematics and computing*, Brooks/Cole, Pacific Grove, CA, 5th ed., 1999.

[51] C. W. Clenshaw and A. R. Curtis, "A method for numerical integration on an automatic computer," *NUMERISCHE MATHEMATIK*, Vol. 2, No. 1, 1960, pp. 197–205.

[52] Bryson, A. and Mills, R., "Linear-quadratic-Gaussian controllers with specified parameter robustness," *JOURNAL OF GUIDANCE CONTROL AND DYNAMICS*, Vol. 21, No. 1, 1998, pp. 11–18.

[53] Coleman, T. and Li, Y., "An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds," *SIAM Journal on Optimization*, Vol. 6, 1996, pp. 418–445.

[54] Coleman, T. and Li, Y., "On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds," *Mathematical Programming*, Vol. 67, No. 2, 1994, pp. 189–224.

[55] Marquardt, D., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *SIAM Journal Applied Math.*, Vol. 11, 1963, pp. 431–441.

[56] Levenberg, K., "A Method for the Solution of Certain Problems in Least-Squares," *Quarterly Applied Math. 2*, 1944, pp. 164–168.

[57] Moré, J., "The Levenberg-Marquardt algorithm: Implementation and theory," *Numerical Analysis*, edited by G. Watson, Vol. 630 of *Lecture Notes in Mathematics*, Springer Berlin / Heidelberg, 1978, pp. 105–116.

[58] Dennis, J.E., J., *Nonlinear Least-Squares, State of the Art in Numerical Analysis*, Academic Press, 1977.

[59] Gelb, A., *Applied Optimal Estimation*, MIT Press, 1974.