JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 36, No. 4, July–August 2013

Polynomial-Chaos-Based Bayesian Approach for State and Parameter Estimations

Reza Madankan,* Puneet Singla,† Tarunraj Singh,‡ and Peter D. Scott§
University at Buffalo, State University of New York, Buffalo, New York 14260

DOI: 10.2514/1.58377

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 probability density function for the input parameters. Later, Galerkin projection is used to obtain a deterministic system of equations for the expansion coefficients. The first proposed approach provides means to update prior expansion coefficients by constraining the polynomial-chaos expansion to satisfy a specified number of posterior moment constraints derived from Bayes's rule. The second proposed approach makes use of the minimum variance formulation to update generalized polynomial-chaos coefficients. The main advantage of the proposed methods is that they not only provide a point estimate for the states and parameters, but they also provide the associated uncertainty estimates along these point estimates. Numerical experiments involving four benchmark problems are considered to illustrate the properties of the proposed methods.

I. Introduction

NUMEROUS fields of science and engineering require the study of the relevant stochastic dynamic system since mathematical models used to represent physical processes and engineering systems have errors and uncertainties associated with them. The major sources of error inherent in any mathematical model prediction consist of errors in model parameters 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 initial-condition 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].

The Kalman filter (KF) is the optimal Bayesian estimator for linear systems with initial condition, model errors, and measurement errors assumed to be Gaussian. However, the performance of the KF can deteriorate appreciably due to model parameter uncertainty [2–4]. 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 (viewed as augmented 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) [12,13] 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 may result 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 the case of a linear system. UKF is one of the approaches that 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. The unscented transformation provides a set of sample points around the mean (known as σ points) that are propagated through the nonlinear functions, from which the mean and covariance of the estimate are then recovered. This process results in a filter that estimates the mean and covariance better than the EKF.

Although both the EKF- and UKF-based filters are widely used for simultaneous state and parameter estimation problems, both methods are based upon a 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 for certain problems based upon the physical model. For example, Gaussian distribution is not an ideal distribution to represent errors in uncertain, positive spring coefficients. This suggests the need for filters that can incorporate the knowledge about non-Gaussian uncertainty. Various researchers have endeavored to exploit knowledge of statistics, dynamic systems, and numerical analysis to develop nonlinear filtering techniques [14–20] that cater to the various classes of state and parameter estimation problems. For low-order nonlinear systems, the particle filter (PF) [18,19] has been gaining increasing attention. However, Daum and Huang in their seminal work [21] discuss that various factors like volume of state space, in which the conditional probability density function (PDF) is nonvanishing; 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 PDFs, etc.); effective dimensionality of the problem; etc.; strongly affect the computational complexity and performance of the particle filter.

For linear systems with parametric uncertainties, the multiple-model estimation [22] method has been very popular. This method assumes the uncertain parameters belong to a discrete set. The

Received 21 March 2012; revision received 5 October 2012; accepted for publication 19 November 2012; published online 5 June 2013.

*Ph.D. Student, Department of Mechanical and Aerospace Engineering; rm93@buffalo.edu.

†Associate Professor, Department of Mechanical and Aerospace Engineering; psingla@buffalo.edu. Senior Member AIAA.

‡Professor, Department of Mechanical and Aerospace Engineering; tsingh@buffalo.edu. Associate Fellow AIAA.

§Professor, Department of Computer Science and Engineering; peter@buffalo.edu.

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. The prior probability values for parameter samples are also updated by making use of Bayes's 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 the number of parameter samples like any sampling algorithm such as the PF [23]. A detailed review on classical approaches applied in online parameter estimation can be found in [24].

All the methods mentioned before 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 the PF encounters expensive computational cost for a large number of samples applied during the estimation process, particularly for high-dimensional models. A useful alternative is to employ spectral representations of uncertain parameters and system states, specifically generalized polynomial-chaos (GPC) expansions for random variables, and stochastic processes.

GPC is an extension of the polynomial-chaos (PC) idea of Wiener [25], which is extensively being used to quantify forward propagation of uncertainty in uncertain dynamic systems. The main principle of the PC 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 uniform distributions, etc.) and transform stochastic equations into deterministic equations in higher-dimensional projection space using Galerkin collocation. 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 [25–27]. The GPC method can efficiently characterize the state uncertainty due to time-invariant random parameters having arbitrary probability distributions.

GPC has been recently used in the Bayesian framework for the parameter estimation problem, also referred to as the inverse problem in the literature [28–32]. All these referenced methods make use of the GPC formulation for the propagation of state or parameter uncertainty through the forward system dynamic model. Pence et al. [28] found a point estimate for the parameter of interest by substituting for GPC expansion in the likelihood function and making use of gradient-based optimization algorithms to solve the resulting maximum likelihood problem. Blanchard et al. [29] proposed a recursive Bayesian approach that makes use of a (suboptimal) EKF to recalculate the PC expansions for the uncertain states and parameters whenever measurement data are available. Marzouk et al. [31] made use of the GPC expansion in conjunction with the Markov chain Monte Carlo (MCMC) to find a maximum posteriori estimate for an uncertain source parameter. More recently, the GPC expansion has been used in a maximum-entropy framework for recursive estimation purposes. Dutta and Bhattacharya [32] developed a nonlinear estimation algorithm based on the combination of the GPC expansion theory, maximum-entropy principle, and higher-order moments updates. However, similar to Schmidt [4], they considered state estimation in the presence of parametric uncertainty. Furthermore, the approximation by Gaussian kernels requires special tuning, which can be cumbersome for many real problems.

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

This paper presents two new recursive approaches to provide estimates for posterior moments of both parameters and system states in the presence of parametric and initial-condition uncertainty by making use of the GPC expansion and the Bayesian framework. The main advantage of the proposed methods is that they not only provide the point estimate (mean) for the state and parameters, but they also provide statistical confidence bounds associated with these estimates described in terms of higher-order posterior moments. Furthermore, these moments have been applied in the construction of posterior coefficients of the GPC expansion for both states and parameters.

The remainder of this paper is structured as follows. In Sec. III, we briefly review the generalized polynomial-chaos theory and its application to the model stochastic differential equations. In Sec. IV, we describe the problem statement and formulation of the estimation process by using Bayes's rule and the minimum variance estimator. Also, detailed formulations of the measurement update process are developed. Next, we illustrate the efficacy of this approach by some numerical examples in Sec. V. Finally, the conclusion and discussion of the results are mentioned.

II. Problem Statement

Consider a general n-dimensional continuous-time dynamic system with uncertain initial conditions and parameters and a discrete-time measurement model, given as

$$\dot{\mathbf{x}}(t, \boldsymbol{\Theta}) = \mathbf{f}(t, \boldsymbol{\Theta}, \mathbf{x}, \mathbf{u}) \tag{1}$$

$$\mathbf{y}_k \triangleq \mathbf{y}(t_k) = \mathbf{h}(\mathbf{x}_k, \boldsymbol{\Theta}) + \boldsymbol{\nu}_k \tag{2}$$

where $\mathbf{x}_k = \mathbf{x}(t_k)$ represents the n-dimensional state vector, the m-dimensional vector $\boldsymbol{\Theta}$ consists of all uncertain time-invariant system and measurement model 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 generally nonlinear function $\mathbf{h}(.)$ captures the measurement model, and the random vector $\boldsymbol{\nu}_k$ denotes the measurement noise with the prescribed distribution $p(\boldsymbol{\nu}_k)$, which is generally assumed to be a zero mean Gaussian PDF. Instead of solving for the point estimates for the state and parameter variables, we are interested in probability distribution for their values. The total uncertainty associated with the state vector $\mathbf{x}(t)$ and parameter vector $\boldsymbol{\Theta}$ is characterized by the PDF $p(t, \mathbf{x}(t), \boldsymbol{\Theta})$, and a nonlinear filtering problem corresponds to finding the a posteriori joint density function for $\mathbf{x}_k$ and $\boldsymbol{\Theta}$ given the measurement data $\mathbf{Y}_k = \{\mathbf{y}_i | i = 1, 2, \ldots, k\}$ [i.e., $p(t, \mathbf{x}(t), \boldsymbol{\Theta} | \mathbf{Y}_k)$] and a prior PDF $p(t_0, \mathbf{x}_0, \boldsymbol{\Theta})$.

As discussed in the last section, several approximate techniques exist in the literature to approximate the posterior state PDF. In the following, we discuss the GPC method for solving the time evolution of the state PDF for systems that include initial-condition and parametric uncertainty.

III. Generalized Polynomial Chaos: Theory and Methodology

This section presents the mathematical details for the GPC methodology to examine the effects of input parameter and initial-condition uncertainty on the forward model outcome. The propagation of uncertainty due to uncertain input parameters and initial conditions can be approximated by a generalization of polynomial-chaos theory. GPC is an extension of the homogenous chaos idea of Wiener [33] 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 uniform distributions, etc. [34]). Galerkin projection is used to generate a system of deterministic differential equations for the expansion coefficients.

A. Linear Systems

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

$$\dot{\mathbf{x}}(t,\boldsymbol{\Theta}) = \mathbf{A}(\boldsymbol{\Theta})\mathbf{x}(t,\boldsymbol{\Theta}) + \mathbf{B}(\boldsymbol{\Theta})\mathbf{u}(t) \tag{3}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{B} \in \mathbb{R}^{n\times p}$. The vector of input signals is $\mathbf{u} \in \mathbb{R}^{p\times 1}$, and $\Theta \in \mathbb{R}^m$ is a vector of uncertain parameters that is a function of the random vector $\boldsymbol{\xi}_\theta = [\xi_{\theta_1}, \xi_{\theta_2}, \cdots, \xi_{\theta_m}]^T \in \mathbb{R}^m$, defined by a PDF $p(\boldsymbol{\xi}_\theta)$ over the support Ω_θ. Similarly, initial conditions $\mathbf{x}(t_0)$ are a function of random vector $\boldsymbol{\xi}_0 = [\xi_{0_1}, \xi_{0_2}, \cdots, \xi_{0_n}]^T$ defined by a PDF $p(\boldsymbol{\xi}_0)$ over the support Ω_0. The random vector $\boldsymbol{\xi}_\theta$ is assumed to be independent of random vector $\boldsymbol{\xi}_0$. Please note that each element of random vector $\boldsymbol{\xi} = [\boldsymbol{\xi}_\theta, \boldsymbol{\xi}_0]^T \in \mathbb{R}^{m+n}$ can be viewed as a component of $m+n$-dimensional stochastic space of random variables. It is assumed that the uncertain state vector $\mathbf{x}(t,\Theta)$ and system parameters A_{ij} and B_{ij} can be written as a linear combination of basis functions $\phi_k(\boldsymbol{\xi})$ that span the stochastic space of random variables $\boldsymbol{\xi} = [\boldsymbol{\xi}_\theta, \boldsymbol{\xi}_0]^T$

$$x_i(t,\boldsymbol{\xi}) = \sum_{k=0}^{N} x_{i_k}(t)\phi_k(\boldsymbol{\xi}) = \mathbf{x}_i^T(t)\boldsymbol{\Phi}(\boldsymbol{\xi}) \Rightarrow \mathbf{x}(t,\boldsymbol{\xi}) = \mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}) \tag{4}$$

$$A_{ij}(\boldsymbol{\xi}) = \sum_{k=0}^{N} a_{ij_k}\phi_k(\boldsymbol{\xi}) = \mathbf{a}_{ij}^T\boldsymbol{\Phi}(\boldsymbol{\xi}) \tag{5}$$

$$B_{ij}(\boldsymbol{\xi}) = \sum_{k=0}^{N} b_{ij_k}\phi_k(\boldsymbol{\xi}) = \mathbf{b}_{ij}^T\boldsymbol{\Phi}(\boldsymbol{\xi}) \tag{6}$$

where $\boldsymbol{\Phi}(.) \in \mathbb{R}^N$ is a vector of polynomial basis functions orthogonal to the PDF $p(\boldsymbol{\xi}) = p(\boldsymbol{\xi}_\theta)p(\boldsymbol{\xi}_0)$, which can be constructed using the Gram–Schmidt orthogonalization process. Table 1 represents different types of polynomial basis functions corresponding to different distributions of random variable $\boldsymbol{\xi}$ [34].

The coefficients $x_{i_k}(t_0)$, a_{ij_k}, and b_{ij_k} are obtained by making use of following normal equations:

$$x_{i_k}(t_0) = \frac{\langle x_i(t_0,\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle}{\langle \phi_k(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle} \tag{7}$$

$$a_{ij_k} = \frac{\langle A_{ij}(\boldsymbol{\Theta}(\boldsymbol{\xi})), \phi_k(\boldsymbol{\xi})\rangle}{\langle \phi_k(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle} \tag{8}$$

$$b_{ij_k} = \frac{\langle B_{ij}(\boldsymbol{\Theta}(\boldsymbol{\xi})), \phi_k(\boldsymbol{\xi})\rangle}{\langle \phi_k(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle} \tag{9}$$

where

$$\langle u(\boldsymbol{\xi}), v(\boldsymbol{\xi})\rangle = \int_{\mathbb{R}^r} u(\boldsymbol{\xi})v(\boldsymbol{\xi})p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}$$

represents the inner product induced by the PDF $p(\boldsymbol{\xi})$.

Note that the total number of terms in the 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 the uncertain parameter $\boldsymbol{\xi}$, which is represented by $m+n$:

$$N = \binom{l+m+n}{m} = \frac{(l+m+n)!}{(m+n)!l!} \tag{10}$$

Now, substitution of Eqs. (4–6) in Eq. (3) leads to

Table 1 Correspondence of polynomial basis functions with their underlying random variables $\boldsymbol{\xi}$

Random variable ξ	Basis polynomials $\phi(.)$	Support
Gaussian	Hermite	$(-\infty, +\infty)$
Gamma	Laguerre	$[0, +\infty]$
Beta	Jacobi	$[a, b]$
Uniform	Legendre	$[a, b]$

$$\mathbf{e}_i(\xi) = \sum_{k=0}^{N} \dot{x}_{i_k}(t)\phi_k(\boldsymbol{\xi}) - \sum_{j=1}^{n}\left(\sum_{k=0}^{N} a_{ij_k}\phi_k(\boldsymbol{\xi})\right)\left(\sum_{k=0}^{N} x_{j_k}(t)\phi_k(\boldsymbol{\xi})\right) - \sum_{j=1}^{p}\left(\sum_{k=0}^{N} b_{ij_k}\phi_k(\boldsymbol{\xi})\right)u_j, \qquad i = 1, 2, \cdots, n \tag{11}$$

Equation (11) represents the error of the approximate GPC solution of Eq. (3), which contains $n(N+1)$ time-varying unknown coefficients $x_{i_k}(t)$. These unknown coefficients can be obtained by using the Galerkin process, i.e., projecting the error of Eq. (3) onto the space of the basis functions $\phi_k(\boldsymbol{\xi})$,

$$\langle e_i(\mathbf{X}_{\text{pc}}, \boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle = 0, \qquad i = 1, 2, \cdots, n, \qquad k = 1, 2, \cdots, N \tag{12}$$

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

$$\dot{\mathbf{x}}_{\text{pc}}(t) = \mathcal{A}\mathbf{x}_{\text{pc}}(t) + \mathcal{B}\mathbf{u}(t) \tag{13}$$

where $\mathbf{x}_{\text{pc}}(t) = \{\mathbf{x}_1^T(T), \mathbf{x}_2^T(T), \cdots, \mathbf{x}_n^T(T)\}$ 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:

$$P_{ij} = \langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi})\rangle, \qquad i, j = 0, 1, 2, \cdots, N \tag{14}$$

$$T_{k_{ij}} = \langle \phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle, \qquad i, j = 0, 1, 2, \cdots, N \tag{15}$$

Then, $\mathcal{A}$ 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 $\mathcal{A}_{ij} \in \mathbb{R}^{(N+1)\times(N+1)}$:

$$\mathcal{A}_{ij} = A_{ij}P, \qquad i, j = 1, 2, \cdots, n \tag{16}$$

if matrix $\mathbf{A}$ is not uncertain; otherwise, it is given by

$$\mathcal{A}_{ij_k} = \mathbf{a}_{ij}^T T_k, \qquad i, j = 1, 2, \cdots, n \tag{17}$$

where $\mathcal{A}_{ij_k}$ represents the kth row of $\mathcal{A}_{ij}$.

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

$$\mathcal{B}_{ij} = Pb_{ij} \qquad i = 1, 2, \cdots, n, \qquad j = 1, 2, \cdots, p \tag{18}$$

if matrix $\mathbf{B}$ is not uncertain; otherwise, it is given by

$$\mathcal{B}_{ij_k} = b_{ij}^T T_k, \qquad i = 1, 2, \cdots, n, \qquad j = 1, 2, \cdots, p \tag{19}$$

where $\mathcal{B}_{ij_k}$ denotes the kth row of $\mathcal{B}_{ij}$.

Equation (4), along with Eq. (13), 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 the uncertain state variable. For example, the first two moments for state vector $\mathbf{x}(t)$ can be written as

$$\mathcal{E}[x_i(t)] = x_{i_1}(t), \qquad i = 1, \cdots, n \tag{20}$$

$$\mathcal{E}[x_i(t)x_j(t)] = \sum_{k=0}^{N} x_{i_k}(t)x_{j_k}(t)\langle \phi_k(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle, \quad i, j = 1, \cdots, n \tag{21}$$

B. Nonlinear Systems with Parametric Uncertainty

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

$$\dot{\mathbf{x}}(t,\boldsymbol{\Theta}) = \mathbf{f}(t,\boldsymbol{\Theta},\mathbf{x},\mathbf{u}), \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{22}$$

where $\mathbf{u}(t)$ is the input to the 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 the uncertain parameter vector $\boldsymbol{\Theta} \in \mathbb{R}^m$ is a function of the random vector $\boldsymbol{\xi}_\theta = [\xi_{\theta_1}, \xi_{\theta_2}, \cdots, \xi_{\theta_m}]^T \in \mathbb{R}^m$ defined by a PDF $p(\boldsymbol{\xi}_\theta)$ over the support Ω_θ. Similarly, initial conditions $\mathbf{x}(t_0)$ are a function of random vector $\boldsymbol{\xi}_0 = [\xi_{0_1}, \xi_{0_2}, \cdots, \xi_{0_n}]^T$ defined by a PDF $p(\boldsymbol{\xi}_0)$ over the support Ω_0. The random vector $\boldsymbol{\xi}_\theta$ is assumed to be independent of random vector $\boldsymbol{\xi}_0$. 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

$$x_i(t,\boldsymbol{\Theta}) = \sum_{k=0}^{N} x_{i_k}(t)\phi_k(\boldsymbol{\xi}) = \mathbf{x}_i^T(t)\boldsymbol{\Phi}(\boldsymbol{\xi}) \Rightarrow \mathbf{x}(t,\boldsymbol{\xi}) = \mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}) \tag{23}$$

$$\theta_i(\boldsymbol{\xi}) = \sum_{k=0}^{N} \theta_{i_k}\phi_k(\boldsymbol{\xi}) = \theta_i^T\boldsymbol{\Phi}(\boldsymbol{\xi}) \Rightarrow \boldsymbol{\Theta}(t,\boldsymbol{\xi}) = \boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}) \tag{24}$$

where $\boldsymbol{\xi}$ is a $m+n$ dimensional vector consisting of all random variables, i.e., $\boldsymbol{\xi} = [\boldsymbol{\xi}_\theta, \boldsymbol{\xi}_0]$. $\mathbf{X}_{\text{pc}}$ and $\boldsymbol{\Theta}_{\text{pc}}$ are matrices composed of coefficients of the GPC expansion for state $\mathbf{x}$ and parameter $\boldsymbol{\Theta}$, respectively. Similar to the linear case, coefficients $x_{i_k}(t_0)$ and θ_{i_k} are obtained by making use of following normal equations:

$$x_{i_k}(t_0) = \frac{\langle x_i(t_0,\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle}{\langle \phi_k(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle} \tag{25}$$

$$\theta_{i_k} = \frac{\langle \theta_i(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle}{\langle \phi_k(\boldsymbol{\xi}), \phi_k(\boldsymbol{\xi})\rangle} \tag{26}$$

Now, substitution of Eqs. (23) and (24) into Eq. (22) leads to

$$\mathbf{e}_i(\mathbf{X}_{\text{pc}},\boldsymbol{\xi}) = \sum_{k=0}^{N} \dot{x}_{i_k}(t)\phi_k(\boldsymbol{\xi}) - \mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}),\boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}),\mathbf{u}), \quad i = 1,2,\cdots,n \tag{27}$$

From Eq. (12), $n(N+1)$ time-varying coefficients x_{i_k} can be obtained using the Galerkin process, i.e., projecting the error captured in Eq. (27) onto the 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 nonpolynomial nonlinearity, such as transcendental or exponential functions, difficulties may arise during the computation of the projection integrals of Eq. (12) . To overcome this issue in the nonlinear case, the polynomial-chaos quadrature technique will be used.

C. Polynomial-Chaos Quadrature

To manage the difficulties in integrating nonpolynomial nonlinearities, Dalbey et al. [35] have proposed a formulation known as the polynomial chaos quadrature (PCQ). PCQ replaces the projection step of the GPC with numerical quadrature. The resulting method can be viewed as a Monte-Carlo-like evaluation of system equations, but with sample points selected by quadrature rules. To illustrate this, consider Eq. (22), which, by substitution of Eqs. (23) and (24), can be written as

$$\sum_{k=0}^{N} \dot{x}_{i_k}(t)\phi_k(\boldsymbol{\xi}) - \mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\xi),\boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}),\mathbf{u}) = 0, \quad i = 1,\cdots,n \tag{28}$$

The projection step of PC yields

$$\sum_{k=0}^{N} \langle \phi_k(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi})\rangle \dot{x}_{i_k} - \langle \mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}),\boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}),\mathbf{u}), \phi_j(\boldsymbol{\xi})\rangle = 0 \quad i = 1,\cdots,n, \qquad j = 0,\cdots,N \tag{29}$$

In the case in which $f(t,\mathbf{x},\boldsymbol{\Theta},\mathbf{u})$ is linear, it is possible to evaluate the projection integrals of Eq. (29) 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 [36] is a suitable choice for most cases. This yields

$$\langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi})\rangle = \int \phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi})p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \simeq \sum_{q=1}^{M} w_q\phi_i(\boldsymbol{\xi}_q)\phi_j(\boldsymbol{\xi}_q) \tag{30}$$

$$\langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi})\phi_k(\boldsymbol{\xi})\rangle = \int \phi_i(\boldsymbol{\xi})\phi_j(\boldsymbol{\xi})\phi_k(\boldsymbol{\xi})p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \simeq \sum_{q=1}^{M} w_q\phi_i(\boldsymbol{\xi}_q)\phi_j(\boldsymbol{\xi}_q)\phi_k(\boldsymbol{\xi}_q) \tag{31}$$

$$\langle \mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}),\boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}),\mathbf{u}), \phi_j(\boldsymbol{\xi})\rangle = \int \mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}),\boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}),\mathbf{u})\phi_j(\boldsymbol{\xi})p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \simeq \sum_{q=1}^{M} w_q\mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\xi_q),\boldsymbol{\Theta}_{\text{pc}}\Phi(\xi_q),\mathbf{u})\phi_j(\xi_q) \tag{32}$$

where M is the number of quadrature points used. Substitution of the aforementioned approximation of the stochastic integral in Eq. (29) and interchanging summation and differentiation leads to

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{q=1}^{M}\sum_{k=0}^{N} w_q\phi_j(\boldsymbol{\xi}_q)\phi_k(\boldsymbol{\xi}_q)x_{i_k} - \sum_{q=1}^{M} w_q\mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}_q),\mathbf{u})\phi_j(\boldsymbol{\xi}_q) = 0 \tag{33}$$

which can be simplified as

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{q=1}^{M} \phi_j(\boldsymbol{\xi}_q)x_i(t,\boldsymbol{\xi}_q)w_q - \sum_{q=1}^{M} w_q\mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}_q),\mathbf{u})\phi_j(\boldsymbol{\xi}_q) = 0 \tag{34}$$

Integrating with respect to time t yields

$$\sum_{q=1}^{M}(x_i(t,\boldsymbol{\xi}_q) - x_i(t_0,\boldsymbol{\xi}_q))\phi_j(\boldsymbol{\xi}_q)w_q - \int_{t_0}^{t}\sum_{q=1}^{M} w_q\mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}_q),\mathbf{u})\phi_j(\boldsymbol{\xi}_q)\,\mathrm{d}t = 0 \tag{35}$$

Interchanging the order of the time integration and quadrature summation leads to

$$\sum_{q=1}^{M}\left\{x_i(t,\boldsymbol{\xi}_q) - x_i(t_0,\boldsymbol{\xi}_q) - \int_{t_0}^{t}\mathbf{f}_i(t,\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}_q),\mathbf{u})\,\mathrm{d}t\right\} \times \phi_j(\boldsymbol{\xi}_q)w_q = 0 \qquad i = 1,\cdots,n \tag{36}$$

Note that the integral expression in Eq. (36) 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 of the output of the system reduces to sampling the chosen input points guided by the quadrature method. Finally, the coefficients of the GPC expansion can be obtained as

$$x_{i_k}(t) = \frac{1}{d_k^2}\sum_{q=1}^{M} \mathcal{X}_i(t_0, t, \boldsymbol{\xi}_q, \mathbf{u})\phi_k(\boldsymbol{\xi}_q)w_q,$$
$$k, j = 0, 1, \cdots, N, \qquad i = 1, 2, \cdots, n \tag{37}$$

where

$$\mathcal{X}_i(t_0, t, \boldsymbol{\xi}_q, \mathbf{u}) = x_i(t_0, \boldsymbol{\xi}_q) + \int_{t_0}^{t} \mathbf{f}_i(t, \mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}_q), \boldsymbol{\Theta}_{\text{pc}}\Phi(\boldsymbol{\xi}_q), \mathbf{u}) \tag{38}$$

$$d_k^2 = \int_{\Omega} \phi_k(\boldsymbol{\xi})\phi_k(\boldsymbol{\xi})p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \tag{39}$$

Hence, the resulting method can be viewed as a Monte-Carlo-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

$$\mathcal{E}[x_i(t)^N] = \int_{\Omega}\left(\int_{t_0}^{t} \dot{x}_i\,\mathrm{d}t\right)^N p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}$$
$$= \int_{\Omega}\left(x_i(t_0, \boldsymbol{\xi}) + \int_{t_0}^{t} \mathbf{f}_i(t, \mathbf{x}, \boldsymbol{\Theta}, \mathbf{u})\,\mathrm{d}t\right)^N p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \quad i = 1, 2, \cdots, n \tag{40}$$

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

$$\mathcal{E}[x_i(t)^N] = \sum_{q}^{M} w_q[\mathcal{X}_i(t_0, t, \boldsymbol{\xi}_q, \mathbf{u})]^N \qquad i = 1, 2, \cdots, n \tag{41}$$

Thus, the output moments can be approximated as a weighted sum of the outputs of simulation runs 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 that is defined by choosing the points optimally in the sense of maximizing the degree of the 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 one-dimensional quadrature points is used to generate quadrature points in general n-dimensional parameter space. As a consequence of this, the number of quadrature points increases exponentially as the number of input parameters increases.

Note that other numerical integration methods like the sparse grid [37] and conjugate unscented transform [38] can also be used to evaluate Eq. (12) . It should be noted that all of these approaches can still suffer from integration error if an insufficient number of samples is 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 the Clenshaw–Curtis quadrature method [39] for numerical integration.

IV. Estimation Process

In the previous section, the GPC theory is presented 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 a 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 to account for its uncertainties and incompleteness. As discussed in Sec. I, 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 section, two different GPC-based approaches have been developed to design finite-dimensional 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 compute accurate prediction between two measurement updates. The first proposed method makes use of Bayes's formula to update the GPC series expansion, while the second method updates the GPC series expansion using the minimum variance technique.

A. Fusion of Measurement Data and Model Estimates

Given a prediction model of Eq. (22), let us consider the sensor model of Eq. (2) . Using the GPC uncertainty evolution as a forecasting tool, the joint PDF of state and parameter can be updated using Bayes's rule on the arrival of a measurement data:

$$p(\boldsymbol{\Theta}, \mathbf{x}|\mathbf{Y}_k) = \frac{p(\boldsymbol{\Theta}, \mathbf{x}|\mathbf{Y}_{k-1})p(\mathbf{y}_k|\boldsymbol{\Theta}, \mathbf{x})}{p(\mathbf{y}_k)} \tag{42}$$

where $\mathbf{Y}_k$ represents the measurement data up to and including time t_k. The joint prior PDF (solution of the GPC approach) of $\mathbf{x}$ and Θ at time t_k given all observations up to time t_{k-1} is $p(\boldsymbol{\Theta}, \mathbf{x}|\mathbf{Y}_{k-1})$, $p(\mathbf{y}_k|\boldsymbol{\Theta}, \mathbf{x})$ is the likelihood that we observe $\mathbf{y}_k$ given $\mathbf{x}$ and $\boldsymbol{\Theta}$ at time t_k, and $p(\boldsymbol{\Theta}, \mathbf{x}|\mathbf{Y}_k)$ represents the joint posterior PDF of $\mathbf{x}$ and $\boldsymbol{\Theta}$ 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:

$$p(\mathbf{y}_k) = \iint p(\boldsymbol{\Theta}, \mathbf{x}|\mathbf{Y}_{k-1})p(\mathbf{y}_k|\boldsymbol{\Theta}, \mathbf{x})\,\mathrm{d}\boldsymbol{\Theta}\,\mathrm{d}\mathbf{x} \tag{43}$$

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

$$\hat{\phi}_k^+ = \mathcal{E}^+[\phi(\boldsymbol{\Theta}, \mathbf{x})] \triangleq \iint \phi(\boldsymbol{\Theta}, \mathbf{x})p(\boldsymbol{\Theta}, \mathbf{x}|\mathbf{Y}_k)\,\mathrm{d}\boldsymbol{\Theta}\,\mathrm{d}\mathbf{x} \tag{44}$$

$$\hat{\phi}_k^- = \mathcal{E}^-[\phi(\boldsymbol{\Theta}, \mathbf{x})] \triangleq \iint \phi(\boldsymbol{\Theta}, \mathbf{x})p(\boldsymbol{\Theta}, \mathbf{x}|\mathbf{Y}_{k-1})\,\mathrm{d}\boldsymbol{\Theta}\,\mathrm{d}\mathbf{x} \tag{45}$$

Now, multiplying Eq. (42) with $\phi(\boldsymbol{\Theta}, \mathbf{x})$ and integrating over $\boldsymbol{\Theta}$ and $\mathbf{x}$, we get

$$\hat{\phi}_k^+ = \frac{\mathcal{E}^-[\phi(\boldsymbol{\Theta}, \mathbf{x})p(\mathbf{y}_k|\boldsymbol{\Theta}, \mathbf{x})]}{p(\mathbf{y}_k)} \tag{46}$$

Note that $\mathbf{y}_k$ is fixed with respect to the expectation operator, and, thus, the right-hand side of Eq. (46) is a function of $\mathbf{y}_k$ only. Notice that Eq. (46) is not an ordinary difference equation, and the evaluation of the right-hand side of Eq. (46) 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 presented in the next section is applicable to any generic likelihood function,

$$p(\mathbf{y}_k|\boldsymbol{\Theta},\mathbf{x}) = \mathcal{N}(\mathbf{y}_k|\mathbf{h}(\mathbf{x}(t),\boldsymbol{\Theta}),\mathbf{R}_k) \triangleq \frac{1}{\sqrt{(2\pi)^k|R_k|}} e^{-(1/2)(\mathbf{y}_k-\mathbf{h}(\mathbf{x}(t),\Theta))^T\mathbf{R}_k^{-1}(\mathbf{y}_k-\mathbf{h}(\mathbf{x}(t),\boldsymbol{\Theta}))} \tag{47}$$

B. 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. (46) in a computationally efficient way. Although the GPC process does not provide us a closed-form expression for the state or parameter PDF, it can be used effectively in computing the expectation integrals. As discussed in the previous section, all moments of random variables $\boldsymbol{\Theta}$ and $\mathbf{x}$ are just a function of their GPC expansion coefficients, i.e., $\boldsymbol{\Theta}_{\text{pc}}$ and $\mathbf{X}_{\text{pc}}$. Hence, one can update the GPC coefficients on the arrival of measurement data based upon Eq. (46). So, if we define $\boldsymbol{\Theta}_{\text{pc}}^-$ and $\mathbf{X}_{\text{pc}}^-$ to be the prior GPC coefficients and $\boldsymbol{\Theta}_{\text{pc}}^+$ and $\mathbf{X}_{\text{pc}}^+$ to be posterior GPC coefficients, then we can evaluate $\hat{\phi}^-(\boldsymbol{\Theta},\mathbf{x})$ and $\hat{\phi}^+(\boldsymbol{\Theta},\mathbf{x})$ as

$$\hat{\phi}_k^- = \hat{\phi}^-(\boldsymbol{\Theta},\mathbf{x}) = \mathcal{E}^-[\phi(\boldsymbol{\Theta},\mathbf{x})] = \int \phi(\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi}),\mathbf{X}_{\text{pc}}^-(t)\Phi(\boldsymbol{\xi}))p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \tag{48}$$

$$\hat{\phi}_k^+ = \hat{\phi}^+(\boldsymbol{\Theta},\mathbf{x}) = \mathcal{E}^+[\phi(\boldsymbol{\Theta},\mathbf{x})] = \int \phi(\boldsymbol{\Theta}_{\text{pc}}^+\Phi(\boldsymbol{\xi}),\mathbf{X}_{\text{pc}}^+(t)\Phi(\boldsymbol{\xi}))p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \tag{49}$$

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

$$M_r(\boldsymbol{\Theta}_{\text{pc}}^-,\mathbf{x}_{\text{pc}}^-,\mathbf{y}_k) = \mathcal{E}^-[\phi(\boldsymbol{\Theta},\mathbf{x})p(\mathbf{y}_k|\boldsymbol{\Theta},\mathbf{x})] = \int \phi(\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi}),\mathbf{X}_{\text{pc}}^-(t)\Phi(\boldsymbol{\xi}))\mathcal{N}(\mathbf{y}_k|\mathbf{h}(\mathbf{X}_{\text{pc}}^-(t) \times \Phi(\boldsymbol{\xi}),\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi})),\mathbf{R}_k)p(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \tag{50}$$

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

$$\mathcal{E}[x_i^+(t)] = x_{i_1}(t), \qquad i = 1,\cdots,n \tag{51}$$

$$\mathcal{E}[x_i^+(t)x_j^+(t)] = \sum_{k=0}^{N} x_{i_k}^+(t)x_{j_k}^+(t)\langle\phi_k(\boldsymbol{\xi}),\phi_k(\boldsymbol{\xi})\rangle, \quad i,j = 1,\cdots,n \tag{52}$$

The main challenge lies in evaluating $M_r(\boldsymbol{\Theta}_{\text{pc}}^-,\mathbf{x}_{\text{pc}}^-,\mathbf{y}_k)$. One can use the quadrature scheme to evaluate Eq. (50):

$$M_r(\boldsymbol{\Theta}_{\text{pc}}^-,\mathbf{x}_{\text{pc}}^-,\mathbf{y}_k) \approx \sum_{q=1}^{N_q} w_q\psi(\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi}_q),\mathbf{X}_{\text{pc}}^-(t)\Phi(\boldsymbol{\xi}_q),\mathbf{y}_k,\mathbf{R}_k) = \sum_{q=1}^{N_q} w_q\psi(\boldsymbol{\Theta}_q^-,\mathbf{x}_q^-,\mathbf{y}_k,\mathbf{R}_k) \tag{53}$$

where

$$\psi(\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi}),\mathbf{X}_{\text{pc}}^-(t)\Phi(\boldsymbol{\xi}),\mathbf{y}_k,\mathbf{R}_k) = \phi(\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi}),\mathbf{X}_{\text{pc}}^-(t)\Phi(\boldsymbol{\xi})) \times\mathcal{N}(\mathbf{y}_k|\mathbf{h}(\mathbf{X}_{\text{pc}}^-(t)\Phi(\boldsymbol{\xi}),\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi})),\mathbf{R}_k) \tag{54}$$

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

$$\phi(\boldsymbol{\Theta},\mathbf{x}) = \boldsymbol{\Theta}_i^{s_1}\mathbf{x}_j^{s_2}, \qquad s_1+s_2 \le N_m \tag{55}$$

Now, substitution of Eqs. (49) and (53) in Eq. (46) leads to

$$g_{s_1,s_2}(\boldsymbol{\Theta}_{\text{pc}}^+,\mathbf{X}_{\text{pc}}^+) = \hat{\phi}^+(\boldsymbol{\Theta}_{\text{pc}}^+,\mathbf{x}_{\text{pc}}^+) - \frac{1}{\alpha}M_r(\boldsymbol{\Theta}_{\text{pc}}^-,\mathbf{x}_{\text{pc}}^-,\mathbf{y}_k), \quad \alpha = p(\mathbf{y}_k), \qquad s_1+s_2 \le N_m \tag{56}$$

where N_c is given as

$$N_c = \sum_{k=1}^{N_m}\frac{(m+n)!}{k!(m+n-k)!}\frac{(N_m)!}{k!(N_m-k)!} \tag{57}$$

and n and m are the dimension of state $\mathbf{x}$ and parameter $\boldsymbol{\Theta}$, respectively. Notice that Eq. (56) is a set of N_c nonlinear coupled equations that defines posterior GPC coefficients $\boldsymbol{\Theta}_{\text{pc}}^+$ and $\mathbf{X}_{\text{pc}}^+$ in terms of prior information that is available from measurement and GPC propagation, to match all joint moments up to order N_m. One can pose the following minimization problem to find a solution for posterior coefficients $\boldsymbol{\Theta}_{\text{pc}}^+$ and $\mathbf{X}_{\text{pc}}^+$:

$$\min_{\boldsymbol{\Theta}_{\text{pc}}^+,\mathbf{X}_{\text{pc}}^+}\left(\sum_{s_1+s_2\le N_m} g_{s_1,s_2}^2(\boldsymbol{\Theta}_{\text{pc}}^+,\mathbf{X}_{\text{pc}}^+)\right) \tag{58}$$

Different algorithms like trust-region-reflective optimization [40,41], the Levenberg–Marquardt optimization [42–44], and the Gauss–Newton approach [44,45] can be used to solve this optimization problem. In this paper, we have used the Levenberg–Marquardt optimization to solve this optimization problem. For a special case for matching just the posterior mean, i.e., $N_m = 1$, we get the following analytical solution for the posterior coefficients:

$$\boldsymbol{\Theta}_{\text{pc}_1}^+ = \kappa_{1,0} \tag{59}$$

$$\mathbf{X}_{\text{pc}_1}^+ = \kappa_{0,1} \tag{60}$$

where $\boldsymbol{\Theta}_{\text{pc}_1}^+$ and $\mathbf{X}_{\text{pc}_1}^+$ represent the first column of $\boldsymbol{\Theta}_{\text{pc}}^+$ and $\mathbf{X}_{\text{pc}}^+$, respectively. Also, $\kappa_{1,0}$ and $\kappa_{0,1}$ are given as

$$\kappa_{1,0} = \sum_{q=1}^{N_q} w_q[\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi})]\mathcal{N}(\mathbf{y}_k|\mathbf{h}(\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi}_q)),\mathbf{R}_k) \tag{61}$$

$$\kappa_{0,1} = \sum_{q=1}^{N_q} w_q[\mathbf{X}_{\text{pc}}^-(t)\Phi(\boldsymbol{\xi})]\mathcal{N}(\mathbf{y}_k|\mathbf{h}(\mathbf{X}_{\text{pc}}(t)\Phi(\boldsymbol{\xi}_q),\boldsymbol{\Theta}_{\text{pc}}^-\Phi(\boldsymbol{\xi}_q)),\mathbf{R}_k) \tag{62}$$

Since the only moment constraint is the 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 the prior value of the rest of the coefficients.

C. GPC-Minimum Variance Estimator

In the previous section, we developed an estimation algorithm to estimate posterior moments and GPC expansion coefficients by making use of Bayes's rule. In this section, we present an alternative development based upon the minimum variance estimator to find an expression for posterior GPC coefficients $\boldsymbol{\Theta}_{\text{pc}}^+$ and $\mathbf{X}_{\text{pc}}^+$. The main advantage of this approach is that it is less computationally demanding than the Bayesian approach described in the last section.

D. Minimum Variance Estimation with A Priori Information

For estimation purposes, we define the concatenated vector $\mathbf{z}$ as

$$\mathbf{z}(t,\boldsymbol{\xi}) = \begin{bmatrix}\mathbf{x}(t,\boldsymbol{\xi})\\ \boldsymbol{\Theta}(\boldsymbol{\xi})\end{bmatrix} \tag{63}$$

Prior and posterior means for both the state and parameter can be written as

$$\hat{\mathbf{z}}_k^- \triangleq \mathcal{E}^-[\mathbf{z}_k] = \begin{bmatrix} \mathbf{X}_{\mathrm{pc}_1}^-(t) \\ \mathbf{\Theta}_{\mathrm{pc}_1}^- \end{bmatrix} \tag{64}$$

$$\hat{\mathbf{z}}_k^+ \triangleq \mathcal{E}^+[\mathbf{z}_k] = \begin{bmatrix} \mathbf{X}_{\mathrm{pc}_1}^+(t) \\ \mathbf{\Theta}_{\mathrm{pc}_1}^+ \end{bmatrix} \tag{65}$$

Similarly, assuming orthonormality of basis functions $\phi_i(\boldsymbol{\xi})$ in the GPC expansion of $\mathbf{x}$ and $\mathbf{\Theta}$, prior and posterior covariance matrices can be written as

$$\Sigma_k^- \triangleq \mathcal{E}^-[(\mathbf{z}_k - \hat{\mathbf{z}}_k^-)(\mathbf{z}_k - \hat{\mathbf{z}}_k^-)^T] = \begin{pmatrix} \sum_{i=1}^N \mathbf{X}_{\mathrm{pc}_i}^{-^2} & \sum_{i=1}^N \mathbf{X}_{\mathrm{pc}_i}^- \mathbf{\Theta}_{\mathrm{pc}_i}^- \\ \sum_{i=1}^N \mathbf{X}_{\mathrm{pc}_i}^- \mathbf{\Theta}_{\mathrm{pc}_i}^- & \sum_{i=1}^N \mathbf{\Theta}_{\mathrm{pc}_i}^{-^2} \end{pmatrix} \tag{66}$$

$$\Sigma_k^+ \triangleq \mathcal{E}^+[(\mathbf{z}_k - \hat{\mathbf{z}}_k^-)(\mathbf{z}_k - \hat{\mathbf{z}}_k^-)^T] = \begin{pmatrix} \sum_{i=1}^N \mathbf{X}_{\mathrm{pc}_i}^{+^2} & \sum_{i=1}^N \mathbf{X}_{\mathrm{pc}_i}^+ \mathbf{\Theta}_{\mathrm{pc}_i}^+ \\ \sum_{i=1}^N \mathbf{X}_{\mathrm{pc}_i}^+ \mathbf{\Theta}_{\mathrm{pc}_i}^+ & \sum_{i=1}^N \mathbf{\Theta}_{\mathrm{pc}_i}^{+^2} \end{pmatrix} \tag{67}$$

where $\mathbf{X}_{\mathrm{pc}_i}^-$ and $\mathbf{\Theta}_{\mathrm{pc}_i}^-$ are the ith column of the PC expansion coefficient matrices $\mathbf{X}_{\mathrm{pc}}^-$ and $\mathbf{\Theta}_{\mathrm{pc}}^-$, respectively. Similarly, $\mathbf{X}_{\mathrm{pc}_i}^+$ and $\mathbf{\Theta}_{\mathrm{pc}_i}^+$ are the ith column of unknown PC expansion coefficient matrices $\mathbf{X}_{\mathrm{pc}}^+$ and $\mathbf{\Theta}_{\mathrm{pc}}^+$, respectively. According to the minimum variance formulation, the posterior mean and covariance can be computed given an estimate of the prior mean and covariance [3]:

$$\hat{\mathbf{z}}_k^+ = \hat{\mathbf{z}}_k^- + \mathbf{K}_k[\tilde{\mathbf{y}}_k - \mathcal{E}^-[\mathbf{h}(\mathbf{x}_k, \mathbf{\Theta})]] \tag{68}$$

$$\Sigma_k^+ = \Sigma_k^- + \mathbf{K}_k \Sigma_{zy} \tag{69}$$

$$\mathbf{K}_k = \Sigma_{zy}^T (\mathbf{P}_{hh}^- + \mathbf{R}_k)^{-1} \tag{70}$$

It should be noted that the minimum variance formulation is valid for any PDF, although it makes use of only mean and covariance information. The sensor output at time t_k is denoted by $\hat{\mathbf{y}}_k$, while the function $\mathbf{h}(\mathbf{x}, \mathbf{\Theta})$ provides a true model between sensor output $\mathbf{y}$ and state $\mathbf{x}$, and parameter $\mathbf{\Theta}$. $\mathbf{R}_k$ denotes the measurement noise error covariance matrix. $\mathbf{K}_k$ is known as the Kalman gain matrix, and matrices Σ_{zy} and Σ_{zz} are defined as

$$\hat{\mathbf{h}}_k^- \triangleq \mathcal{E}^-[\mathbf{h}(\mathbf{x}_k, \mathbf{\Theta})] = \sum_{q=1}^M w_q \underbrace{\mathbf{h}(\mathbf{x}_k(\boldsymbol{\xi}_q))}_{\mathbf{h}_q} \tag{71}$$

$$\begin{aligned} \Sigma_{zy} &\triangleq \mathcal{E}^-[(\mathbf{z}_k - \hat{\mathbf{z}}_k)(\mathbf{h}(\mathbf{x}_k) - \hat{\mathbf{h}}_k^-)^T] \\ &= \sum_{q=1}^M w_q (\mathbf{z}_k(\boldsymbol{\xi}_q) - \hat{\mathbf{z}}_k^-)(\mathbf{h}_q - \hat{\mathbf{h}}_k^-)^T \end{aligned} \tag{72}$$

$$\begin{aligned} \Sigma_{hh}^- &\triangleq \mathcal{E}^-[(\mathbf{h}(\mathbf{x}_k) - \hat{\mathbf{h}}_k^-)(\mathbf{h}(\mathbf{x}_k) - \hat{\mathbf{h}}_k^-)^T] \\ &= \sum_{q=1}^M w_q (\mathbf{h}_q - \hat{\mathbf{h}}_k^-)(\mathbf{h}_q - \hat{\mathbf{h}}_k^-)^T \end{aligned} \tag{73}$$

Notice that Eqs. (65) and (68) provide a closed-form solution for $\mathbf{X}_{\mathrm{pc}_1}^+$ and $\mathbf{\Theta}_{\mathrm{pc}_1}^+$, while one can solve for the rest of the posterior coefficients by making use of Eqs. (67) and (69).

V. Numerical Examples

In the previous section, we have developed two algorithms based upon the GPC expansion for state and parameter estimations. Here, we consider four different numerical experiments to demonstrate the performance of these methods. We also employ the EKF and bootstrap PF algorithms to compare the performance of the proposed methodology. All the simulations are performed in the MATLAB environment and on a dual-core desktop computer with a 2.13 GHz Intel Core 2 CPU.

A. First-Order System

As the first example, we consider the forced first-order system

$$\dot{x} + Kx = U_{\mathrm{in}}, \qquad x(0) = 0 \tag{74}$$

$$y(t_k) = x(t_k) + \nu_k \tag{75}$$

where $U_{\mathrm{in}} = 2e^{-t/10}\sin(2t)$ and the prior uncertainty in K is assumed to be uniformly distributed over the interval [0.5, 1.5]. For simulation purposes, the measurement data are assumed to be available at a sampling frequency of 1 Hz. A random sample of K is taken from the prior distribution to generate the noise-free measurement data. The noise-free measurement data are corrupted with a Gaussian white noise of zero mean and variance being 0.05. It should be mentioned that this simulation is performed for different realizations of measurement noise and values of parameter K. The results presented here correspond to the true value of K being 1.3659 ($K_{\mathrm{act}} = 1.3659$). To represent uncertainty in state and parameter, a ninth-order GPC expansion is considered, and the total simulation time interval is assumed to be 10 s. The initial GPC expansion for K and $x(0)$ can be written as

$$x(0, \boldsymbol{\xi}) = \sum_{k=0}^{9} x_k(0)\phi_k(\boldsymbol{\xi}) \qquad x_k(0) = 0 \tag{76}$$

$$K(\boldsymbol{\xi}) = \sum_{i=0}^{9} k_i \phi_i(\boldsymbol{\xi}) \quad k_0 = 1, \quad k_2 = 0.5 \ \text{ and } \ k_i = 0 \tag{77}$$

where $\phi_k(\boldsymbol{\xi})$ are Legendre polynomials that correspond to the uniform distribution of parameter K. Using the procedure outlined in Sec. III, Eq. (74) can be converted into the following deterministic form:

$$\mathcal{M}\dot{\mathbf{X}}_{\mathrm{pc}}(t) + \mathcal{K}\mathbf{X}_{\mathrm{pc}} = \begin{pmatrix} 2e^{-t/10}\sin(2t) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{78}$$

where

$$\mathcal{M}_{i+1,j+1} = \langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle = \frac{1}{2i+1}\delta_{ij}, \qquad i, j = 0, 1, \cdots, N \tag{79}$$

$$\begin{aligned} \mathcal{K}_{i+1,j+1} &= \langle \phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle + 0.5\langle \phi_1(\boldsymbol{\xi})\phi_i(\boldsymbol{\xi}), \phi_j(\boldsymbol{\xi}) \rangle, \\ & i, j = 0, 1, \cdots, N \end{aligned} \tag{80}$$

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. (78) is given by

Table 2 Relative error of moments of state x with respect to 100,000 Monte Carlo runs at $t = 2$ s

Number of quadrature points	Mean, %	Second central moment, %	Third 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
1000 MC simulations	1.2310	3.7051	7.8948

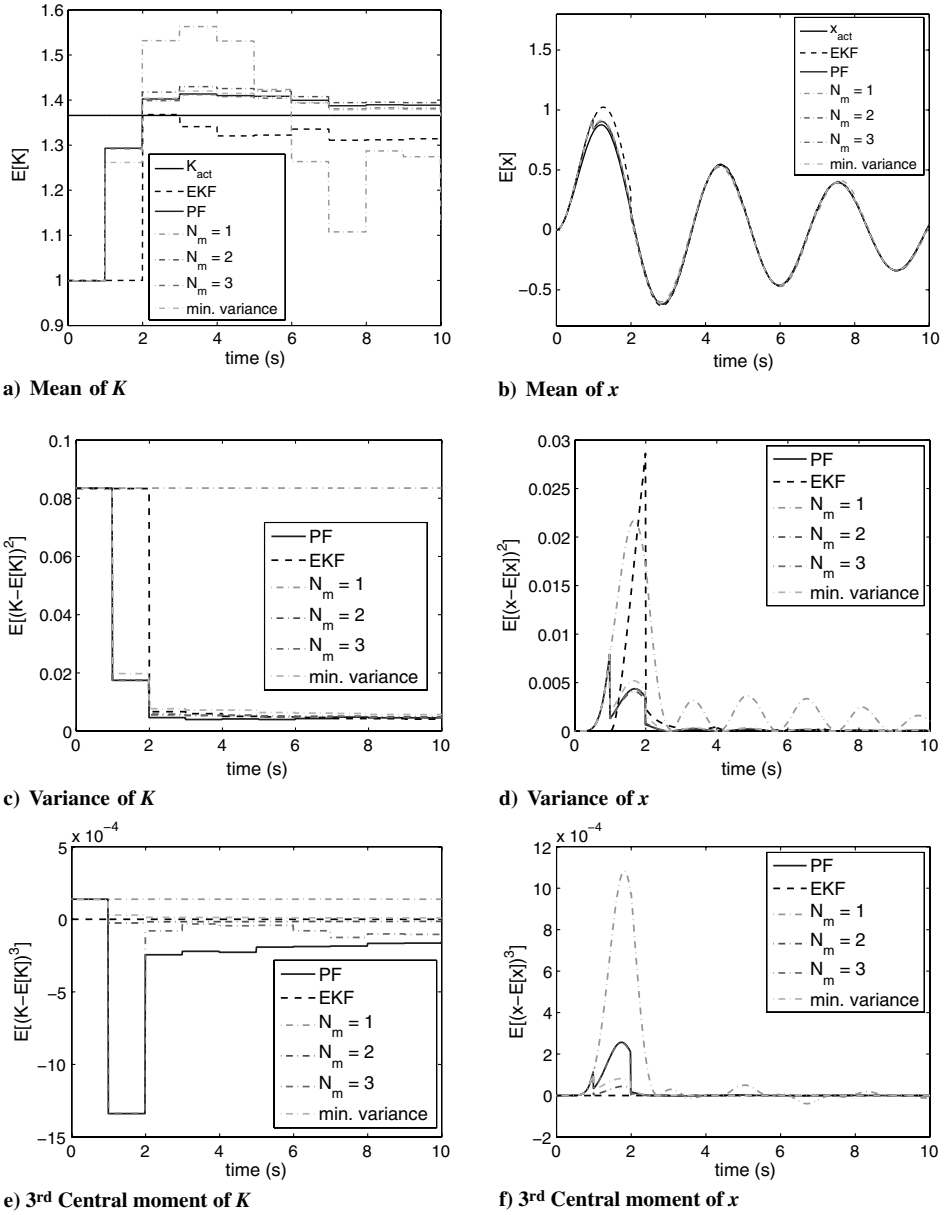
$$x_i(0) = 0 \qquad i = 0, \cdots, N \tag{81}$$

where N is the number of terms used in the GPC expansion of x. The solution of this system of ordinary differential equations yields the coefficients of GPC expansion of $x(t)$, which can be used in Eq. (23) to construct the solution of Eq. (74).

After studying the convergence of first three central moments vs number of Monte Carlo (MC) runs [46], 100,000 MC runs are considered to be the reference truth for this example to verify the efficacy of the GPC method in forward propagation. Table 2 shows the relative error in approximating the first three central moments using the PCQ framework for x at $t = 2$ s. It should be noted that one needs only four quadrature points or model runs according to the PCQ formulation to capture the first three moments with less than 1% error while 1000 MC runs result in an order-of-magnitude-higher error when compared against 100,000 MC runs. These results clearly show the efficacy of the GPC framework in accurately propagating the parameter uncertainty through the dynamical system.

The mean estimates for parameter K and state x by using different estimation algorithms (PF, EKF, GPC-based minimum variance estimator, and the GPC–Bayes method for different moment matching constraints (different values of N_m)) have been shown in Figs. 1a and 1b, respectively. As expected, the GPC–Bayes method results in more accurate results as we increase N_m and assume the PF approximated posterior mean to be the reference truth. Also, when $N_m = 2$, the GPC–Bayes and GPC-based minimum variance estimators perform very similarly in finding the 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.

Figures 1c and 1d show the posterior variance for parameter K and state x corresponding to different filters, respectively. As expected, the GPC–Bayes approach with $N_m = 1$ cannot capture the posterior variance for parameter K and state x. However, the performance of



a) Mean of K b) Mean of x

c) Variance of K d) Variance of x

e) 3rd Central moment of K f) 3rd Central moment of x

Fig. 1 Posterior central moments for parameter K and state x for the first-order system.

the GPC–Bayes method improves significantly in capturing the posterior variance as compared to the PF by increasing N_m, i.e., the 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.

Figures 1e and 1f 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 the 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 the third central moment for both K and x.

Tables 3 and 4 show the root mean square error (RMSE) 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 RMSE. 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 the importance function in the measurement update part. As expected, the GPC–Bayes method results in less error in estimation of the 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 the estimation of the first two central moments for x.

Finally, Table 5 represents the processor time corresponding to the implementation of different algorithms. Clearly, the EKF performs faster than all the other methods. However, due to nonlinearities involved in the augmented dynamical system, it results in poor estimation results. The GPC-based minimum variance estimator performs slower than the EKF, but it results in much more accurate estimates than the EKF. From these results, it is clear that the GPC–Bayes method with low values of N_m is much faster than the PF. As expected, the computational cost corresponding to the GPC–Bayes method increases as one increases the number of matching moment constraints (N_m). This increase in computational cost can be attributed to solving the optimization problem at every measurement update step.

In summary for this example, it is clear that the proposed methods perform 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 the higher-order posterior moments. The poor performance of the EKF algorithm can be attributed to the strong nonlinearity involved due to the simultaneous state and parameter estimation problem.

Table 3 RMSE in the first three posterior central moments for parameter K assuming PF with 100,000 particles to be the reference truth

N_m	Mean	Second central moment	Third 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
Minimum variance	3.6602e − 001	6.9057e − 002	1.4929e − 002
EKF	4.5070e + 000	5.6481e − 001	1.4602e − 002

Table 4 RMSE in the first three posterior central moments for state x assuming a PF with 100,000 particles to be the reference truth

N_m	Mean	Second central moment	Third 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
Minimum variance	1.3997e − 001	8.5417e − 003	1.2629e − 003
EKF	1.4636e + 000	3.6966e − 002	1.9292e − 003

Table 5 Processor time (in seconds) required for different estimation approaches for the first-order system

N_m	EKF	PF	Minimum variance	GPC–Bayes
1	— —	— —	— —	464.46
2	0.8	2541.4	112.09	681.52
3	— —	— —	— —	8162.8

B. Duffing Oscillator

We next consider the Duffing oscillator

$$\ddot{x} + \eta\dot{x} + \alpha x + \beta x^3 = \sin(3t) \tag{82}$$

$$\mathbf{y}(t_k) = \begin{bmatrix} x(t_k) \\ \dot{x}(t_k) \end{bmatrix} + \nu_k \tag{83}$$

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

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

The initial states are assumed to be normally distributed:

$$x(0) = \mathcal{N}(x_0| - 1, 0.5^2), \qquad \dot{x}(0) = \mathcal{N}(\dot{x}_0| - 1, 0.5^2)$$

Hence, $\psi_k(\boldsymbol{\xi})$ are chosen to be Hermite polynomials to describe the Gaussian distribution of states. As well, the fourth-order GPC expansion is considered to analyze the effect of the initial-condition uncertainty. To corroborate the efficacy of the PCQ approach to capture the evolution of the statistics of the states of Eq. (82), a relative error in the Frobenius norm of the difference between different moments of states with respect to 100,000 Monte Carlo runs at $t = 2$ s is evaluated. Table 6 shows that the relative error decreases as the number of quadrature points increases. It is clear that one can obtain a better approximation for three central moments using only 16 quadrature points, relative to the 1000 MC runs.

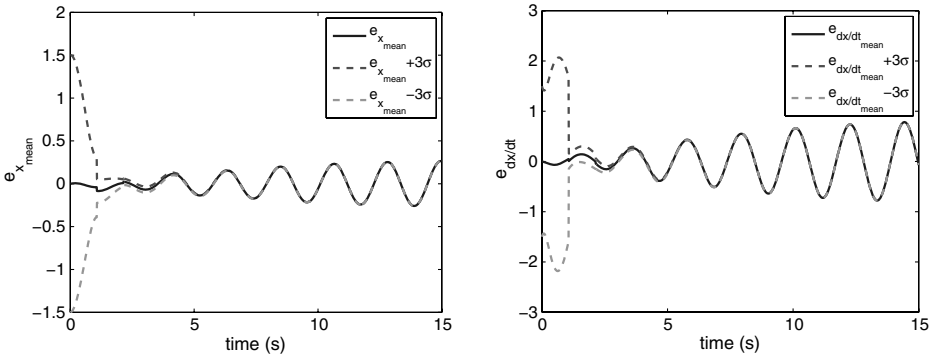
To verify the efficiency of our method, we compared the performance of the proposed methods with the EKF and PF results. The measurement data are assumed to be available at a sampling frequency of 1 Hz. A random sample of initial conditions is taken from initial-condition distribution to generate the noise-free measurement data. The noise-free measurement data are then corrupted with a Gaussian white noise of zero mean and variance being

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

where σ is assumed to be 0.05 in our simulations. Figures 2a and 2b illustrate the state estimation error for x and $\dot{x}$ by using the EKF

Table 6 Relative error in the Frobenius norm of the difference between moments of the states and 100,000 Monte Carlo runs at $t = 2$ s

Number of quadrature points	Mean, %	Second central moment, %	Third central moment, %
1^2	4.5526	100	100
2^2	0.3217	20.3050	98.3149
3^2	0.0329	3.9559	28.5219
4^2	0.0537	0.5202	2.5084
1000 MC simulations	0.1199	6.0715	99.2219



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

Fig. 2 Error and 3σ bounds for the EKF approximated posterior mean for the Duffing oscillator.

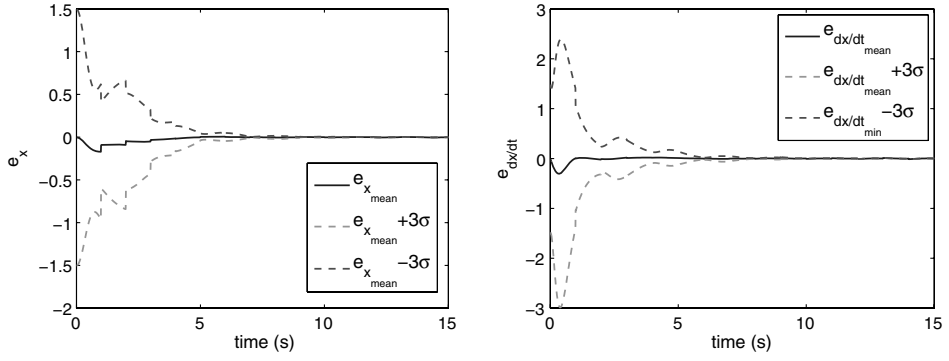
method, respectively. The solid line represents the difference between the true value and its mean estimate, while the dashed lines show the $\pm 3\sigma$ bounds. From these plots, it is clear that the state estimation error increases significantly with time although it is always bounded by 3σ bounds. The poor performance of the EKF can be attributed to strong nonlinearities and sparse data resulting from sampling at 1 Hz.

The state estimation error for x and $\dot{x}$ by using the particle filter has been shown in Figs. 3a and 3b, respectively. The solid line represents the difference between the true value and its mean estimate, while the dashed lines show the minimum and the maximum bounds. These plots show that the state estimation error decreases with time, while using the PF.

Furthermore, Fig. 4 shows the error in the state estimates along with its 3σ bounds using the GPC-based minimum variance estimator. Once again, the estimation error along with 3σ bounds converge to zero over the time, which can be again attributed to the posterior density function converging to a delta function as the number of measurements increases.

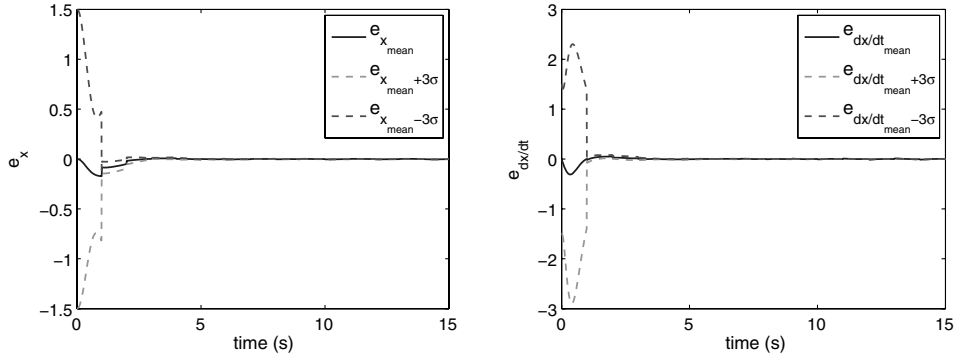
Figure 5 shows the error in the state estimates using the GPC–Bayes method for various values of N_m. The solid line represents the difference between the true value and its mean estimate, while the dashed lines represent the minimum and maximum bounds on the estimation errors. It is clear that the estimation error and corresponding 3σ bounds for the estimation error converge to zero over time. This is due to the fact that the posterior density function finally converges to a Dirac-delta function around the truth, which is expected as the number of measurements increases over time. Also, it should be noticed that 3σ bounds become tighter and tighter as one increases the number of matching moment constraints, i.e., N_m. From these results, it is clear that the proposed methods perform very well in not only estimating the posterior mean but the posterior density function also.

To summarize, the RMSE over time between the mean estimate of states and their true value has been shown in Table 7. As this table represents, the GPC–Bayes, PF, and GPC-based minimum variance method perform very well in the estimation of both states x and $\dot{x}$, while the EKF results in high error between the mean estimate and the actual value of the states. It is clear from Table 7 that, by increasing the number of matching moment constraints (N_m) in the GPC–Bayes method, the error in the estimation of the states decreases.



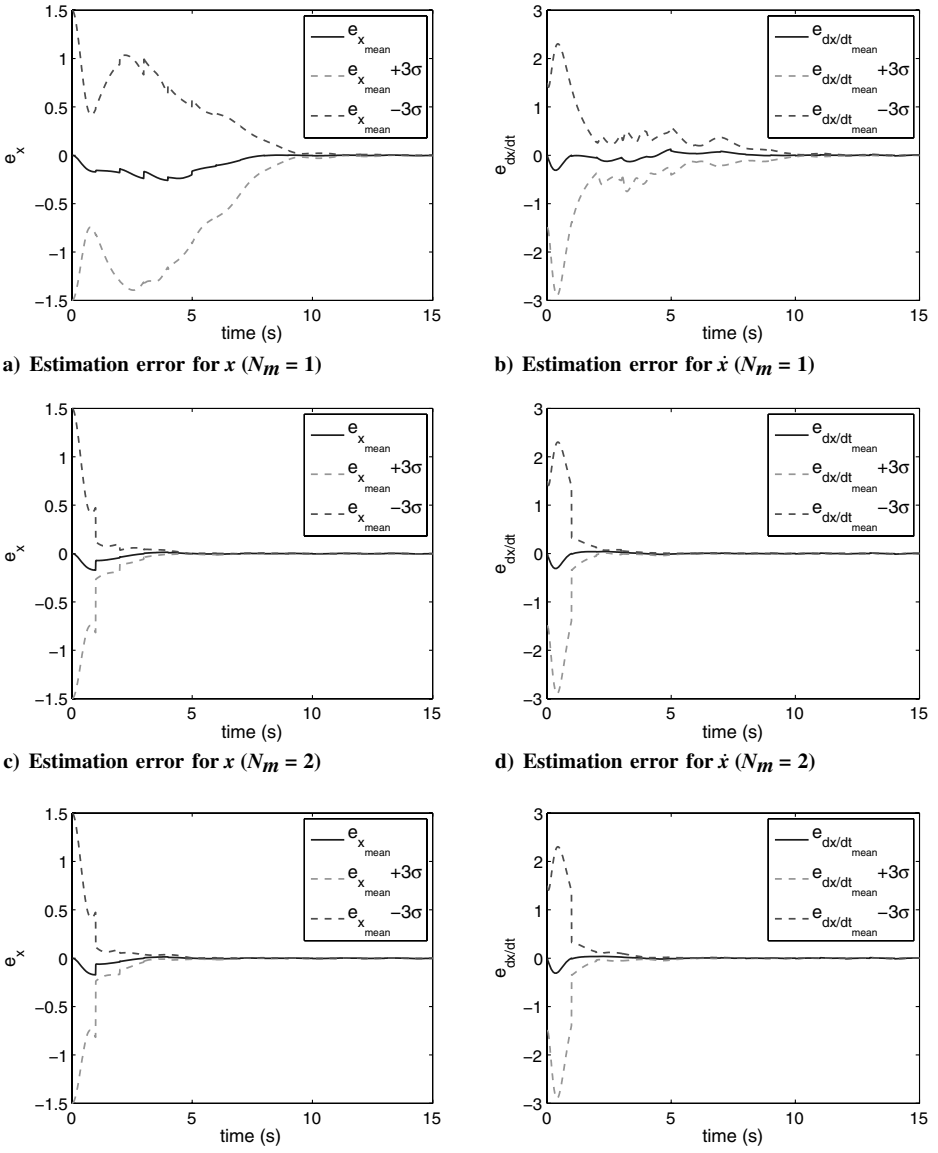
a) $x(t)$ **b)** $\dot{x}(t)$

Fig. 3 Error and 3σ bounds for PF approximated posterior mean for the Duffing oscillator.



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

Fig. 4 Error and 3σ bounds for the minimum variance approximated posterior mean for the Duffing oscillator.



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 $\dot{x}$ ($N_m = 3$)

Fig. 5 Estimation error and 3σ bounds for the GPC–Bayes approximated posterior mean for the Duffing oscillator.

Finally, Table 8 shows the processor time required for various algorithms. As expected, the EKF implementation is the fastest, but it results in erroneous estimates. Both the GPC-based minimum variance estimator and the GPC–Bayes filter require less computational effort than the PF implementation, while both methods perform well in estimating the system states.

C. Falling Body Problem

We now consider the benchmark problem of the falling body with the nonlinear observation model as originally described by Athans et al. [47]:

$$\dot{x}_1 = -x_2 \quad \dot{x}_2 = -x_2^2 p e^{-\lambda x_1} \tag{84}$$

$$\mathbf{y}(t_k) = \sqrt{M^2 + x_1^2} + \nu_k \tag{85}$$

where x_1 represents the altitude of the body in feet and x_2 represents the downward velocity in feet/second. The ballisitc coefficient p is defined as follows:

$$p \triangleq C_D \frac{A\rho_0}{2m}$$

Table 7 RMSE in the mean estimate of states x and $\dot{x}$ while using different estimation methods

Method		e_x	$e_{\dot{x}}$
EKF		0.1307	0.3858
PF		0.0408	0.0508
Minimum variance		0.0359	0.0527
GPC–Bayes	$N_m = 1$	0.1173	0.0695
	$N_m = 2$	0.0347	0.0531
	$N_m = 3$	0.0336	0.0528

Table 8 Processor time (seconds) required for the different estimation approaches

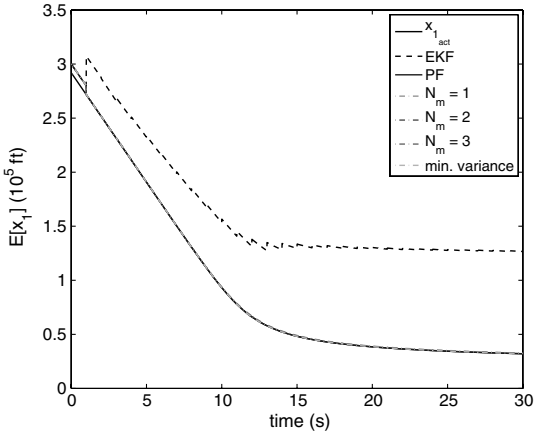
Method		Processor time, s
EKF		1.2399
PF		26,411
Minimum variance		53.4602
GPC–Bayes	$N_m = 1$	65.8251
	$N_m = 2$	64.6385
	$N_m = 3$	108.9398

Table 9 Relative error in the Frobenius norm of the difference between the moments of the states and 100,000 Monte Carlo runs at $t = 2$ s

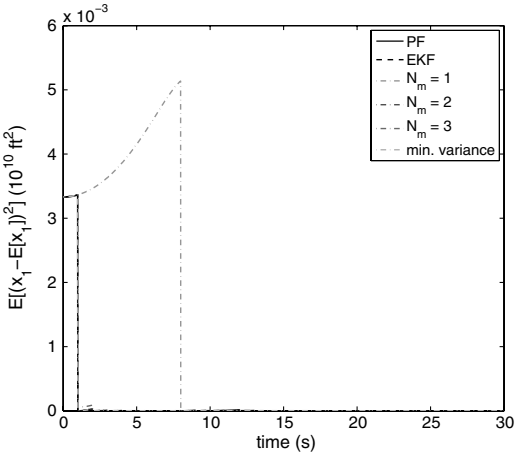
Number of quadrature points	Mean, %	Second central moment, %	Third central moment, %
1^3	3.2802	100	100
2^3	2.43e − 007	0.0020	0.0979
3^3	2.189e − 008	0.0020	0.0004
1000 MC simulations	0.0331	2.3890	94.5458

where C_D, A, ρ_0, and m represent the drag coefficient, reference area for drag evaluation, mass density of the atmosphere, and mass of the falling object, respectively. The measurements are assumed to be the range of the body as measured from a radar location located at a horizontal distance $M = 100,000$ ft away from the falling body. For simulation purposes, the initial states and parameter are assumed to be uniformly distributed as follows:

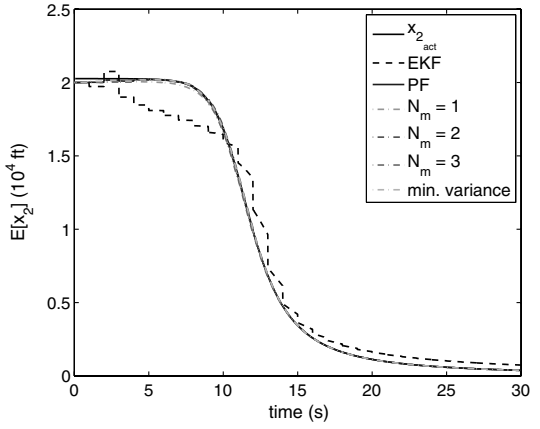
$$x_1(0) = U(2.9 \times 10^5, 3.1 \times 10^5),$$
$$x_2(0) = U(1.9 \times 10^4, 2.1 \times 10^4),$$
$$p = U(5 \times 10^{-4}, 1.5 \times 10^{-3})$$



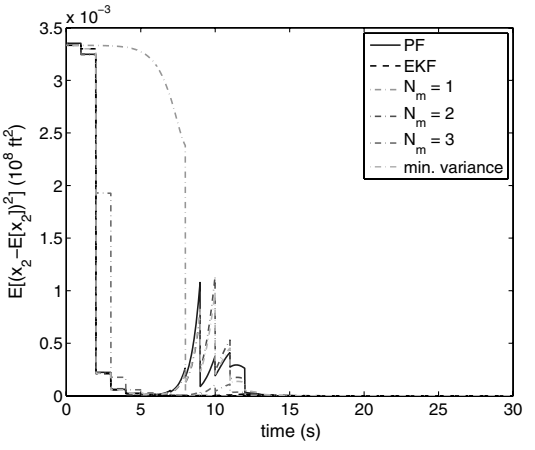
a) Mean of x_1



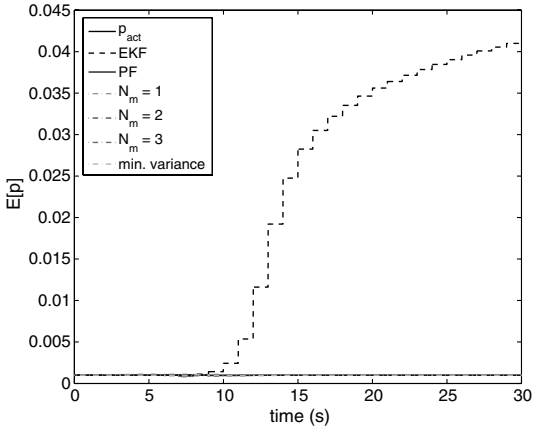
b) Variance of x_1



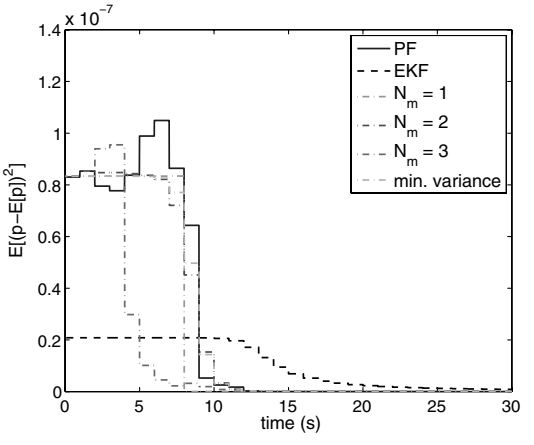
c) Mean of x_2



d) Variance of x_2



e) Mean of p



f) Variance of p

Fig. 6 Posterior expected value of the states and parameter for the falling body model.

Please note that, due to large initial values of altitude and velocity, the effect of gravity is negligible and has not been considered in the model [47]. Legendre polynomials $\phi_k(\boldsymbol{\xi})$ are chosen to describe the uniform distribution of states. A third-order GPC expansion is considered to analyze the effect of the initial-condition uncertainty. To compare efficacy of the PCQ approach to capture the evolution of the statistics of the states of Eq. (84), a relative error in the Frobenius norm of the difference between different moments of the first two states with respect to 5×10^4 Monte Carlo runs at $t = 2$ s is evaluated. Table 9 shows that the relative error decreases as the number of quadrature points increases. It is clear that one can obtain a better approximation for three central moments using only eight quadrature points, relative to the 1000 MC runs.

We compared the performance of the proposed methods with the EKF and PF results. The measurement data are assumed to be available at a sampling frequency of 1 Hz. The following initial conditions are randomly taken from initial conditions and parameter distribution to generate the noise-free measurement data:

$$x_{1_{\text{act}}}(0) = 2.9195 \times 10^5, \qquad x_{2_{\text{act}}}(0) = 2.0265 \times 10^4,$$
$$p_{\text{act}} = 9.8859 \times 10^{-4}$$

The noise-free measurement data are then corrupted with a Gaussian white noise of zero mean and variance $R = 100^2$.

Figure 6 represents the mean and variance estimates corresponding to various filters for the states. Also, the Frobenius norm of the RMSE over time in approximating the first three central moments for states are listed in Table 10. The PF-estimated posterior central moment is considered to be the reference truth while computing the RMSE. As expected, the accuracy of the GPC–Bayes method in capturing the first three central moments improves as the number of matching moment constraints is increased. The performance of the GPC–Bayes filter in approximating the central moments degrades over time, especially for parameter p, which can be attributed to the finite-order GPC approximation. The EKF performs the worst among all the filters as it even fails to capture the posterior mean. Similar to the previous examples, the GPC-based minimum variance filter performs significantly better than the EKF in capturing the first two central moments. For detailed comparison between GPC–Bayes, PF, and minimum variance method, please see Fig. 7.

Table 11 shows the processor time required for the implementation of various algorithms in the MATLAB simulation environment. Once again, the EKF is faster than all other approaches but results in erroneous estimates. As expected, the computational cost for the GPC–Bayes method increases as one increases the number of matching moment constraints (N_m). However, both the GPC-based minimum variance estimator and the GPC–Bayes method are faster than the PF and also provide good estimates for the posterior moments as illustrated in Table 10.

Table 10 Frobenius norm of RMSE in the first three scaled posterior central moments for states x_1, x_2 and parameter p

N_m	Mean	Second central moment	Third central moment
1	6.5620e − 001	1.3327e − 001	1.2258e − 004
2	2.0491e − 001	4.9733e − 003	6.5058e − 005
3	5.8985e − 001	1.8427e − 002	1.1854e − 004
Minimum variance	1.2564e − 001	4.2480e − 003	7.3951e − 005
EKF	4.1950e + 001	8.8058e − 003	7.0741e − 005

Table 11 Computational time (seconds) required for different estimation approaches for the falling body problem

N_m	EKF	PF	Minimum variance	GPC–Bayes
1	1.2240	8.6144e + 003	4.0782e + 001	1.1376e + 002
2				3.5331e + 002
3				2.4081e + 003

D. Hovering Helicopter Model

As the last example, we examine the efficiency of the proposed approach on a hovering helicopter model given by

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} = \begin{pmatrix} p_1 & p_2 & -g & 0 \\ 1.26 & -1.765 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} - \begin{pmatrix} 0.086 \\ -7.408 \\ 0 \\ 0 \end{pmatrix} K_{lqr} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \tag{86}$$

$$\mathbf{y}(t_k) = \begin{bmatrix} x_1(t_k) \\ x_2(t_k) \\ x_3(t_k) \\ x_4(t_k) \end{bmatrix} + \nu_k \tag{87}$$

where K_{lqr} and the initial conditions are given as

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

The state vector $\mathbf{x}$ describes the horizontal velocity x_1 in feet/second, the pitch angle of the fuselage x_2 in centiradians, its derivative x_3 in centiradians/s, and perturbation x_4 in feet from a ground point reference. Coefficient g corresponds to the acceleration due to gravity given by 0.322. Parameters p_1 and p_2 are assumed to be uniformly distributed over the intervals $[-0.2, 0]$ and $[0, 0.2]$, respectively. For simulation purposes, measurement data are assumed to be available at a sampling frequency of 1 Hz. A random sample of unknown parameters is taken from their prior distributions to generate the noise-free measurement data. The noise-free measurement data are corrupted with a Gaussian white noise of zero mean and standard deviation being 0.15 times an identity matrix. To represent uncertainty in the state and parameter, a seventh-order GPC expansion is considered, and the total simulation time interval is assumed to be 10 s.

To verify the efficiency of the PCQ framework in the simulation of the forward propagation of Eq. (86), a relative error in the Frobenius norm of the difference between different moments of all states with respect to 100,000 Monte Carlo runs at $t = 2$ s is evaluated. Table 12 shows that the relative error decreases as the number of quadrature points increases. It is clear that only nine quadrature points result in better approximation for three central moments, in comparison with 1000 MC runs.

Figures 8 and 9 show the posterior mean estimate corresponding to various filters for the parameters and states, respectively. Furthermore, the Frobenius norm of the RMSE over time in approximating the first three central moments for parameters and states are listed in Tables 13 and 14. We assume the PF estimated posterior central moment to be the reference truth to compute the RMSE. As expected, the accuracy of the GPC–Bayes method in capturing the first three central moments improves as the number of matching moment constraints is increased. The performance of the GPC–Bayes filter in approximating the third central moment degrades over time, which can be attributed to the finite-order GPC approximation. The EKF

Table 12 Relative error in the Frobenius norm of the difference between the moments of the states and 100,000 Monte Carlo runs at $t = 2$ s

Number of quadrature points	Mean, %	Second central moment, %	Third central moment, %
1^2	2.9776	100	100
2^2	0.0121	2.5408	69.1695
3^2	0.0013	0.0268	0.8993
4^2	7.2606×10^{-6}	1.5195×10^{-4}	0.0082
1000 MC simulations	19.3794	57.6130	86.2600

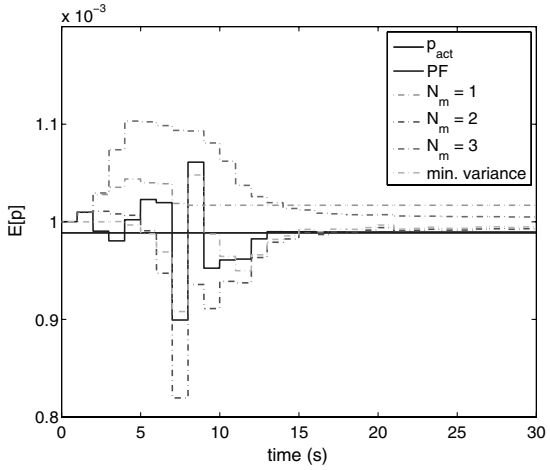
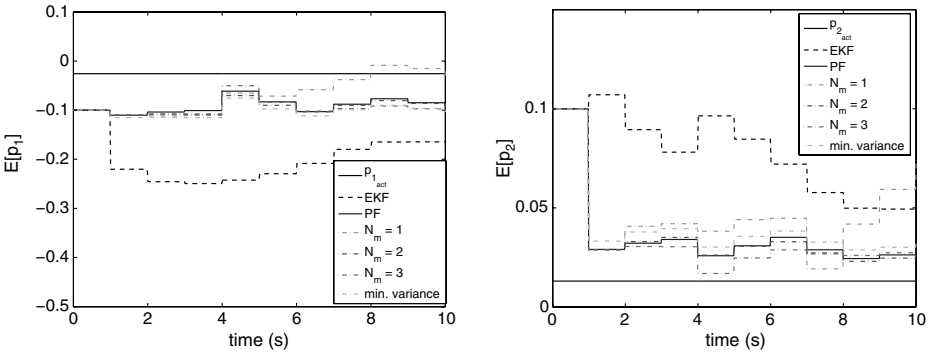


Fig. 7 Closer view for posterior expected value of parameter p for the falling body model, using the PF, minimum variance, and GPC–Bayes method ($N_m = 1,2,3$).

performs the worst among all 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.

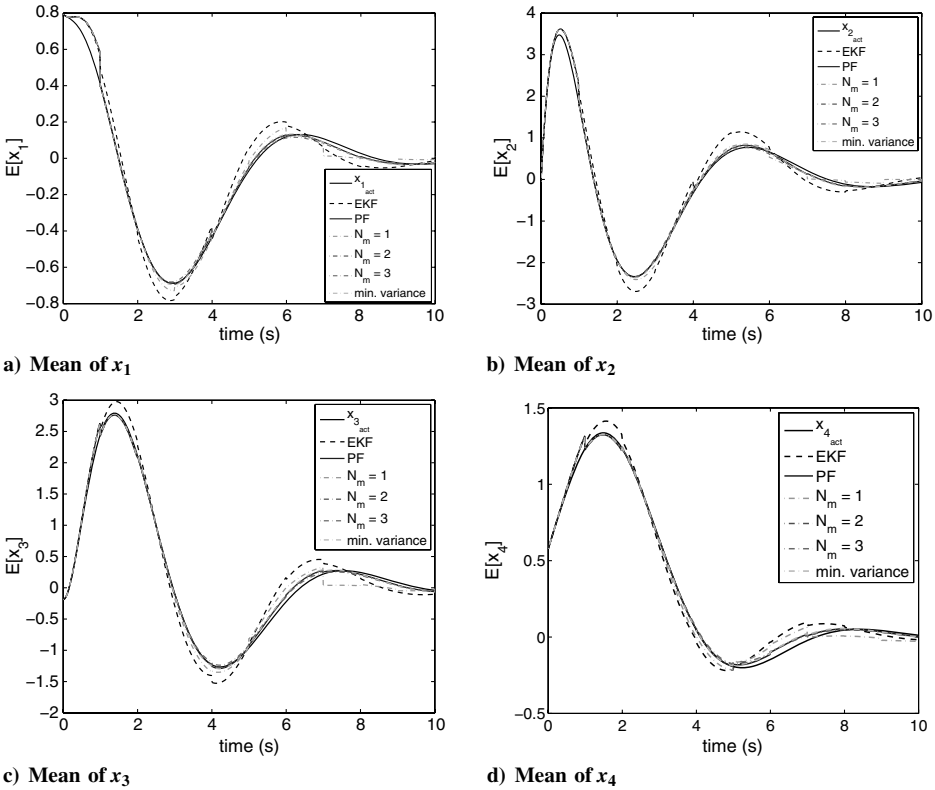
Finally, Table 15 shows the processor time required for the implementation of different algorithms. Once again, the EKF is the fastest algorithm, but it performs very poorly in estimating the states and parameters. The GPC-based minimum variance estimator performs slower than the EKF, but it results in much more accurate estimates than the EKF. Furthermore, it is clear that the GPC–Bayes method with low values of N_m is much faster than the PF. As expected, the computational cost corresponding to the GPC–Bayes method increases as one increases the number of matching moment constraints (N_m). This increase in computational cost can be attributed to solving the optimization problem at every measurement update step which can be reduced with the help of more efficient optimization routines.

As we noticed in the last example, 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



a) Mean of $\mathbf{p_1}$ **b) Mean of $\mathbf{p_2}$**

Fig. 8 Posterior expected value of parameters for the hovering helicopter model.



a) Mean of x_1 **b) Mean of x_2**

c) Mean of x_3 **d) Mean of x_4**

Fig. 9 Posterior expected value of the states (x_1, x_2, x_3, and x_4) for the hovering helicopter model.

Table 13 Frobenius norm of RMSE in the first three posterior central moments for parameters p_1 and p_2

N_m	Mean	Second central moment	Third central moment
1	1.2788e − 003	9.7578e − 005	1.1902e − 006
2	1.5801e − 004	1.3676e − 005	1.2375e − 006
3	2.8708e − 004	2.5960e − 005	2.9830e − 006
Minimum variance	3.8147e − 004	1.3044e − 005	1.2020e − 006
EKF	4.1991e − 002	9.2605e − 005	1.1919e − 006

Table 14 Frobenius norm of RMSE in the first three posterior central moments for states x_1, x_2, x_3, and x_4

N_m	Mean	Second central moment	Third central moment
1	3.9557e − 003	1.1085e − 002	6.3803e − 003
2	4.7443e − 004	7.4849e − 005	2.0311e − 005
3	1.2053e − 003	8.1340e − 005	1.8096e − 005
Minimum variance	4.8027e − 004	5.3403e − 005	9.2885e − 006
EKF	2.8215e − 003	7.8116e − 004	4.3769e − 005

Table 15 Processor time (seconds) required for different estimation approaches applied to the hovering helicopter example

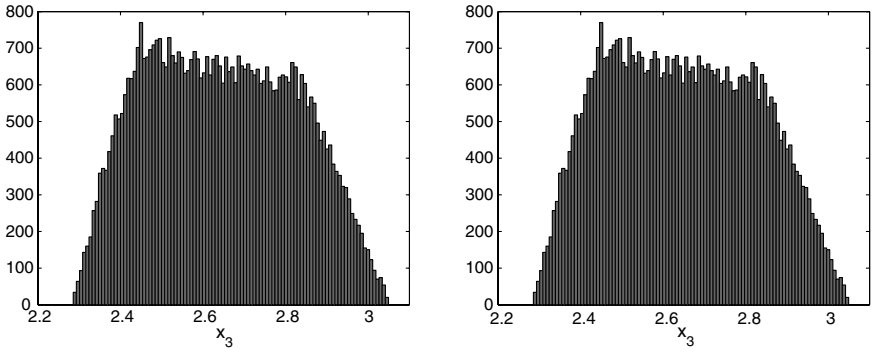
N_m	EKF	PF	Minimum variance	GPC–Bayes
1	1.1454	8.8284e + 003	2.5649e + 001	2.6441e + 002
2				6.0449e + 002
3				3.4292e + 004

Table 16 First three posterior central moments of state x_3 at $t = 1$ s by using the PF and GPC–Bayes method ($N_m = 3$)

Method	Mean	Second central moment	Third central moment
PF	2.4302	2.341e − 003	1.1603e − 005
GPC–Bayes ($N_m = 3$)	2.4300	3.503e − 003	1.4304e − 005

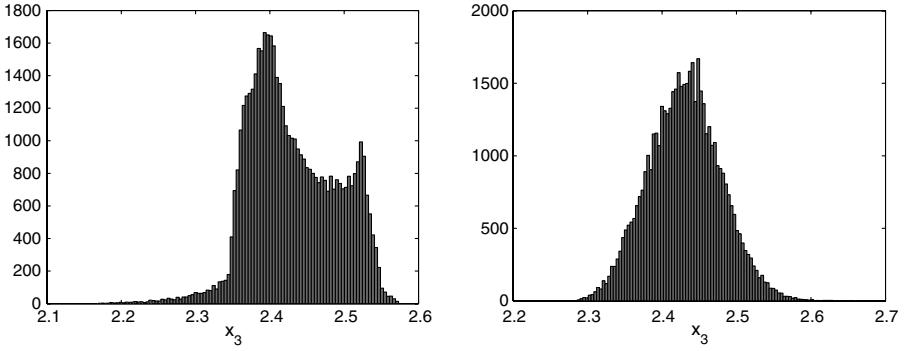
more detail, let us reconsider one of the states, e.g., x_3. As Fig. 10 represents, given identical realizations of applied random variables, both the GPC and Monte Carlo methods results in identical empirical distribution for x_3 at $t = 1$ s.

Notice that all three posterior central moments for x_3 approximated by the GPC–Bayes method match well with the PF approximation up to the first measurement update, i.e., $t = 1$ s for $N_m = 3$. To make this point more clear, the first three posterior central moments for x_3 at $t = 1$ s are listed in Table 16. Furthermore, Fig. 11 shows histograms corresponding to the posterior distribution for x_3 after the first measurement update at $t = 1$ s for the GPC–Bayes ($N_m = 3$) and the PF. From Fig. 11, it is clear that, even though the first three central moments of x_3 are the same at $t = 1$ s, the posterior distributions approximated by using the PF and the GPC–Bayes method are bit different due to a mismatch in the 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 the histogram plots of Fig. 12 and the moment data in Table 17. Ideally, one can overcome this error by increasing the number of moment matching constraints and the order of the GPC expansion, which leads to a higher computational load. In practice, one needs to compromise between the computational load and accuracy in approximating the higher-order moments.



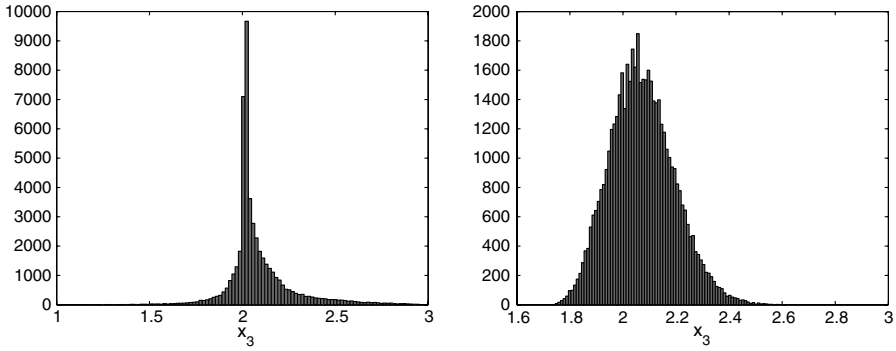
a) GPC-Bayes ($N_m = 3$) b) Particle filter

Fig. 10 Histograms of state x_3 before the measurement update at $t = 1$ s.



a) GPC-Bayes ($N_m = 3$) b) Particle filter

Fig. 11 Histograms of state x_3 after the measurement update at $t = 1$ s.



a) GPC-Bayes ($N_m = 3$) **b) Particle filter**

Fig. 12 Histograms of state x_3 before the measurement update at $t = 2$ s.

Table 17 First three central moments of state x_3 before the measurement update at $t = 2$ s by using the PF and GPC–Bayes method ($N_m = 3$)

Method	Mean	Second central moment	Third central moment
PF	2.0725	1.543e − 002	6.2095e − 004
GPC–Bayes ($N_m = 3$)	2.0822	1.478e − 002	1.4224e − 004

VI. Conclusions

In this research, two new recursive approaches have been developed to provide accurate estimates for the posterior moments of both parameters and system states while making use of the generalized polynomial-chaos framework for the uncertainty propagation. The main advantage of the proposed methods is that they not only provide a point estimate for the state and parameters, but they also permit the calculation of statistical confidence bounds associated with these estimates.

The numerical results show that the proposed methodologies perform better than the extended Kalman filter in capturing the posterior moments for both the state and parameter. Furthermore, it is demonstrated that one can converge to the particle filter estimates for not only the posterior mean but also the higher-order moments by increasing the number of matching moment constraints in the generalized polynomial-chaos–Bayes method. The generalized polynomial-chaos-based minimum variance method demonstrated consistently valid estimates of the posterior mean and variance for both the state and parameters. The numerical example results show that the processor time associated with the polynomial-chaos-based minimum variance estimator and the generalized polynomial-chaos–Bayes method with one or two matching moment constraints is much lower than that associated with the particle filter, while consistently providing accurate estimates for the 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 better spectral content of the posterior density function.

An open research issue is to associate the error in approximating the moments with the order of the polynomial-chaos expansion. This kind of error analysis can help one in selecting the order of the polynomial-chaos expansion to match the desired order of moments. However, this analysis is difficult due to the absence of any closure in the moment space.

Acknowledgment

This material is based upon work supported by the National Science Foundation under award number CMMI-1054759 and Air Force Office of Scientific Research (AFOSR) grant number FA9550-11-1-0012.

References

[1] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, Vol. 82, No. 1, 1960, pp. 35–45. doi:10.1115/1.3662552

[2] Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, Dover, New York, 1970, pp. 281–286.

[3] Gelb, A., *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974, pp. 182–190, 203–214.

[4] Schmidt, S. F., "Application of State-Space Methods to Navigation Problems," *Advances in 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, July 2001, pp. 998–1013. doi:10.1109/9.935054

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

[7] Simon, D., *Optimal State Estimation: Kalman, H_∞ and Nonlinear Approaches*, Wiley, Hoboken, NJ, 2006, pp. 361–369.

[8] Petersen, I. R., and Savkin, A. V., *Robust Kalman Filtering for Signals and Systems with Large Uncertainties*, Birkhäuser, 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. doi:10.1016/j.automatica.2004.02.014

[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*, Dover, New York, 1979, pp. 193–222, 277–297.

[12] Julier, S. J., and Uhlmann, J. K., "New Extension of the Kalman Filter to Nonlinear Systems," *AeroSense'97*, International Society for Optics and Photonics, 1997, pp. 182–193.

[13] Julier, S., and Uhlmann, J., "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, Vol. 92, IEEE Publ., Piscataway, NJ, 2004, pp. 401–422.

[14] Daum, F. E., "A New Nonlinear Filtering Formula for Discrete Time Measurements," *24th IEEE Conference on Decision and Control*, Vol. 24, IEEE Publ., Piscataway, NJ, Dec. 1985, pp. 1957–1958.

[15] Daum, F. E., "A New Nonlinear Filtering Formula Non-Gaussian Discrete Time Measurements," *25th IEEE Conference on Decision and Control*, Vol. 25, IEEE Publ., Piscataway, NJ, Dec. 1986, pp. 1030–1031.

[16] Daum, F. E., "Exact Finite Dimensional Nonlinear Filters," *24th IEEE Conference on Decision and Control*, Vol. 24, IEEE Publ., Piscataway, NJ, Dec. 1985, pp. 1938–1945.

[17] Ito, K., and Xiong, K., "Gaussian Filters For Nonlinear Filtering Problems," *IEEE Transactions on Automatic Control*, Vol. 45, No. 5, 2000, pp. 910–927. doi:10.1109/9.855552

[18] Sloan, I. H., and Woniakowski, H., "When Are Quasi-Monte Carlo Algorithms Efficient for High Dimensional Integrals?" *Journal of Complexity*, Vol. 14, No. 1, 1998, pp. 1–33. doi:10.1006/jcom.1997.0463

[19] Djuric, P., and Goodsill, S., "Guest Editorial Special Issue on Monte Carlo Methods for Statistical Signal Processing," *IEEE Transactions on Signal Processing*, Vol. 50, No. 2, Feb. 2002, p. 173. doi:10.1109/TSP.2002.978373

[20] Terejanu, G., Singla, P., Singh, T., and Scott, P. D., "Uncertainty Propagation for Nonlinear Dynamic Systems Using Gaussian Mixture Models," *Journal of Guidance, Control, and Dynamics*, Vol. 31, No. 6, 2008, pp. 1623–1633. doi:10.2514/1.36247

[21] Daum, F., and Huang, J., "Curse of Dimensionality and Particle Filters," *Proceedings of IEEE Aerospace Conference*, Vol. 4, IEEE Publ., Piscataway, NJ, March, 2003, pp. 1979–1993.

[22] Anderson, B. D. O., and Moore, J. B., *Optimal Filtering*, Dover, New York, 1979, pp. 277–297.

[23] 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. doi:10.1109/78.978374

[24] 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.

[25] Wiener, N., "The Homogeneous Chaos," *American Journal of Mathematics*, Vol. 60, No. 4, 1938, pp. 897–936. doi:10.2307/2371268

[26] 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. doi:10.1137/S1064827501387826

[27] Ghanem, R. G., and Spanos, P. D., *Stochastic Finite Elements: A Spectral Approach*, Springer–Verlag, New York, 1991, pp. 42–47.

[28] Pence, B. L., Fathy, H. K., and Stein, J. L., "A Maximum Likelihood Approach to Recursive Polynomial Chaos Parameter Estimation," *Proceedings of the American Control Conference (ACC)*, IEEE, 2010, pp. 2144–2151.

[29] Blanchard, E., Sandu, A., and Sandu, C., "Parameter Estimation Method Using an Extended Kalman Filter," *Proceedings of the Joint North America, Asia-Pacific ISTVS Conference and Annual Meeting of Japanese Society for Terramechanics*, International Society for Terrain-Vehicle Systems, 2007, pp. 23–26.

[30] Li, J., and Xiu, D., "A Generalized Polynomial Chaos Based Ensemble Kalman Filter with High Accuracy," *Journal of Computational Physics*, Vol. 228, No. 15, 2009, pp. 5454–5469. doi:10.1016/j.jcp.2009.04.029

[31] 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, 2007, pp. 560–586. doi:10.1016/j.jcp.2006.10.010

[32] Dutta, P., and Bhattacharya, R., "Nonlinear Estimation with Polynomial Chaos and Higher Order Moment Updates," *Proceedings of the American Control Conference (ACC)*, IEEE, 2010, pp. 3142–3147.

[33] Wiener, N., "The Homogeneous Chaos," *American Journal of Mathematics*, Vol. 60, No. 4, 1938, pp. 897–936. doi:10.2307/2371268

[34] 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, p. 619. doi:10.1137/S1064827501387826

[35] 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," *Journal of Geophysical Research*, Vol. 113, B5, 2008, pp. 1–16. doi:10.1029/2006JB004471

[36] Abramowitz, M., and Stegun, I. A., *Handbook of Mathematical Functions (with Formulas, Graphs, and Mathematical Tables), Integration*, Dover, New York, 1972, Chap. 25.4.

[37] Gerstner, T., and Griebel, M., "Numerical Integration Using Sparse Grids," *Numerical Algorithms*, Vol. 18, No. 3, 1998, pp. 209–232. doi:10.1023/A:1019129717644

[38] Adurthi, N., Singla, P., and Singh, T., "The Conjugate Unscented Transform—An Approach to Evaluate Multi-Dimensional Expectation Integrals," *Proceedings of the American Control Conference (ACC)*, IEEE, 2012, pp. 5556–5561.

[39] Clenshaw, C. W., and Curtis, A. R., "A Method for Numerical Integration on an Automatic Computer," *Numerische Mathematik*, Vol. 2, No. 1, 1960, pp. 197–205. doi:10.1007/BF01386223

[40] Coleman, T., and Li, Y., "An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds," *SIAM Journal on Optimization*, Vol. 6, No. 2, 1996, pp. 418–445. doi:10.1137/0806023

[41] Coleman, T. F., and Li, Y., "On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds," *Mathematical Programming*, Vol. 67, No. 1, 1994, pp. 189–224.

[42] Marquardt, D., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *SIAM Journal on Applied Mathematics*, Vol. 11, No. 2, 1963, pp. 431–441. doi:10.1137/0111030

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

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

[45] Bjorck, A., *Numerical Methods for Least Squares Problems*, Vol. 51, Society for Industrial Mathematics, 1996, pp. 342–348.

[46] Madankan, R., *Polynomial Chaos Based Method for State and Parameter Estimation*, M.S. Thesis, Univ. at Buffalo, State University of New York, Buffalo, NY, 2011.

[47] Athans, M., Wishner, R., and Bertolini, A., "Suboptimal State Estimation for Continuous-Time Nonlinear Systems from Discrete Noisy Measurements," *IEEE Transactions on Automatic Control*, Vol. 13, No. 5, 1968, pp. 504–514. doi:10.1109/TAC.1968.1098986