A Polynomial Chaos-Based Nonlinear Bayesian Approach for Estimating State and Parameter Probability Distribution Functions

Vinay A. Bavdekar and Ali Mesbah

Abstract— **Various systems engineering and control applications require the knowledge of the complete probability distribution function (pdf) of system states and parameters. This work presents a nonlinear Bayesian estimation approach that uses the histogram filter algorithm to construct the posterior pdfs of the state variables and uncertain parameters based on histograms of the prior and likelihood pdfs. To address the computational challenges associated with the Bayesian estimation algorithms in obtaining the posterior pdfs, the generalized polynomial chaos framework is used to enable efficient propagation of the time-invariant probabilistic system uncertainties with arbitrary distributions. The proposed estimation approach is demonstrated on a benchmark continuous bioreactor, and its performance and computational requirements are compared to those of a sequential importance resampling particle filter.**

I. Introduction

State and parameter estimation plays a crucial role in many engineering applications including process monitoring, model-based design, optimization, and control. State estimation algorithms are mostly based on the recursive Bayesian estimation framework, which determines the posterior probability distribution function (pdf) of states/parameters conditioned on system measurements (e.g., [1], [2], [3]). For nonlinear systems, however, there exists no closed-form solution to the Bayes' rule [2]. Thus, nonlinear state estimation algorithms primarily aim at developing and improving approximate solutions to the Bayes' rule. State estimation for nonlinear systems is further compounded by model uncertainty, as well as exogenous disturbances that are ubiquitous in complex dynamical systems.

Various nonlinear Bayesian state estimation algorithms such as the ensemble Kalman filter (EnKF) [4] and particle filters (PFs) [2] use sample-based techniques to account for the system uncertainties in estimating the posterior pdf of the states. For linear systems, histogram filter (HF) is proposed to construct the histograms of the prior and likelihood using randomly drawn samples of the states, disturbances, and noise [5]. The prior and likelihood histograms are used in the Bayes' rule to construct the histogram of the posterior pdf of the states. The sample-based estimation algorithms mainly use Monte Carlo (MC) techniques to draw random samples from the known pdfs of uncertainties. However, the computational burden associated with MC-based Bayesian state estimation algorithms can be prohibitive for nonlinear systems with large state dimension. This shortcoming arises from the need to repeatedly solve the system model for every uncertainty realization.

This work uses the generalized polynomial chaos (gPC) framework [6] to address the recursive Bayesian estimation problem for nonlinear systems with probabilistic time-invariant uncertainties and (time-varying) system disturbances. The gPC framework allows for systematically accounting for the effect of arbitrary, time-invariant uncertainties in model parameters and initial conditions. In the gPC framework, the stochastic state variables are expressed as an expansion of orthogonal polynomial basis functions, the coefficients of which yield the statistical moments of the stochastic states. The polynomial chaos expansions can also be used as a computationally efficient surrogate for the original nonlinear model to perform MC simulations. The gPC framework has been used as an efficient uncertainty propagation tool in various applications such as stochastic MPC [7], [8], active fault diagnosis [9], optimal experiment design [10], and state estimation [11], [12], [13], [14], [15].

Konda et al. [14] proposed two gPC-based uncertainty propagation approaches for state estimation of linear systems subject to time-invariant parametric uncertainties and Gaussian disturbances. In the first approach, the mean and covariance of the (Gaussian) states are expanded with respect to parametric uncertainties using the gPC framework. In the second approach, the states are mapped onto the space of coefficients of the polynomial chaos expansions, which are evaluated for different realizations of the Gaussian disturbances. The mean and variance of states are then used for estimating the statistics of the posterior state pdfs. For nonlinear systems with uncertainties in model parameters and initial conditions, [11] and [12] proposed, respectively, a gPC-based ensemble Kalman filter (gPC-EnKF) and a gPC-based extended Kalman filter (gPC-EKF). The prediction and measurement update steps of the gPC-based filters are defined in terms of the coefficients of PC expansion, which are used to compute the moments of the state variables. In [13], the gPC framework is used in conjunction with the Gaussian mixture approximation to compute the posterior pdf of states. Madankan et al. [15] used the Bayes' rule to compute the moments of the posterior pdf of states, based on which the coefficients of the PC expansion are updated. For nonlinear systems subject to stochastic disturbances and parametric uncertainties, Madankan et. al. [16] proposed the use of a conjugate unscented transform for computing the first two moments of the posterior pdf of states. The work assumes that the states have symmetric pdfs around the mean, which can be a restrictive assumption for nonlinear

Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA, 94720.
{bavdekar, mesbah}@berkeley.edu

systems. Except for [14] and [16], the majority of the gPC-based nonlinear state estimation algorithms only account for the effect of time-invariant parametric and initial condition uncertainties, while disregarding the effect of disturbances. Further, most of the existing gPC-based state estimation algorithms only determine the point estimates of the states as well as their moments, instead of characterizing the entire posterior pdf of states.

This work presents a Bayesian state estimation algorithm for nonlinear systems with arbitrary probabilistic uncertainties and disturbances. The proposed algorithm adopts the gPC framework for propagation of the time-invariant probabilistic uncertainties, and uses the principles of the HF algorithm for constructing the posterior pdf of the uncertain states and parameters. The key features of the proposed algorithm are as follows: i) it accounts for time-invariant probabilistic uncertainties as well as system disturbances, ii) it utilizes the computational advantages offered by the gPC framework, in which MC simulations are performed through algebraic operations, to eliminate the need for repeatedly solving the system model as in MC-based estimation techniques such as the PF, and iii) it approximates the posterior pdf of states (and parameters), instead of merely estimating the moments of the posterior distribution. The proposed algorithm is demonstrated on a benchmark bioreactor simulation case study [17]. The performance and computational requirements of the proposed algorithm are compared to those of an MC-based sequential importance resampling (SIR) particle filter [2].

Notation.

$\mathbb{N} = \{1, 2, \ldots\}$ denotes the set of natural numbers, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. $\mathsf{P}(\cdot)$ denotes the pdf of a stochastic variable. $\mathsf{P}(\cdot|z)$ denotes the conditional pdf, conditioned on z. $\mathcal{N}(\mu, \Sigma)$ denotes a Gaussian distribution with mean μ and covariance Σ. $\mathcal{N}(x; \mu, \Sigma)$ denotes the probability value of x, given the distribution $\mathcal{N}(\mu, \Sigma)$. $\mathbf{E}[\cdot]$ denotes the expected value of a stochastic variable.

II. Problem Formulation

Consider a discrete-time nonlinear system

$$x_k = f(x_{k-1}, u_{k-1}, \theta_0) + w_{k-1}, \tag{1a}$$

$$y_k = h(x_k) + v_k \tag{1b}$$

where $k \in \mathbb{N}_0$ is the time index; $x_k \in \mathbb{R}^n$ denotes the system states; $u_k \in \mathbb{R}^{n_u}$ denotes the system inputs; $\theta_0 \in \mathbb{R}^p$ denotes the true system parameters; $y_k \in \mathbb{R}^r\,(r \leq n)$ denotes the measured system outputs; $f(\cdot)$ denotes the nonlinear state dynamics; $h(\cdot)$ denotes the measurement function; $w_k \in \mathbb{R}^n$ is zero-mean random system disturbances with a known pdf $\mathsf{P}(w)$; and $v_k \in \mathbb{R}^r$ is zero-mean measurement noise with a known pdf $\mathsf{P}(v)$. It is assumed that the function $f(\cdot)$ is of polynomial form or can be transformed into a polynomial-in-states form [18]. For this assumption to hold, the sufficient conditions are that $f(\cdot)$ is analytic with respect to x_k and separable with respect to x_k, u_k, and θ [18]. The function $h(\cdot)$ is also assumed to be analytic and separable with respect to x_k.

The model used to describe the dynamics of (1) is subject to probabilistic uncertainties arising from imperfect knowledge of the model parameters and initial states. The model parameters $\{\theta_i\}_{i=1}^p$ are independently distributed with pdf $\mathsf{P}(\theta_i),\ i = 1, 2, \ldots, p$. The initial states are also uncertain with a known distribution $x_0 \sim \mathsf{P}(x_0)$. Define a probability space $\{\Omega, \mathcal{F}, \mathsf{P}\}$ on the sample space Ω, σ-algebra $\mathcal{F}$, and the probability measure P on Ω. The time-invariant uncertainties $\begin{bmatrix} x_0^\top & \theta^\top \end{bmatrix} \in \mathbb{R}^{n_\xi},\ (n_\xi \leq n + p)$ are defined in terms of the standard random variables $\xi \in \mathbb{R}^{n_\xi}$ such that $\xi_j \in L^2\{\Omega, \mathcal{F}, \mathsf{P}\}$, where $L^2\{\cdot\}$ is the Hilbert space of ξ_j and $\mathbf{E}\left[\xi_j^2\right] < \infty$. The elements $\{\xi_j\}_{j=1}^{j=n_\xi}$ are assumed to be independently distributed with known pdfs $\mathsf{P}(\xi_j)$.

The objective of this work is to estimate the posterior pdf of the states and uncertain parameters of system (1) using the Bayes' rule [2]. Under the assumption that (1) is a Markov process, the Bayes' rule is expressed as

$$\mathsf{P}(x_k|y_k) = \frac{\mathsf{P}(y_k|x_k)\,\mathsf{P}(x_k|y_{k-1})}{\mathsf{P}(y_k|y_{k-1})}, \tag{2}$$

where $\mathsf{P}(x_k|y_{k-1})$ denotes the prior pdf of the states based on the measurements y_{k-1}; $\mathsf{P}(y_k|x_k)$ denotes the likelihood, that is, the probability of the measurements given a particular value of the current states; $\mathsf{P}(y_k|y_{k-1})$ is the evidence, a normalizing constant that is the marginal of the measurements; and $\mathsf{P}(x_k|y_k)$ denotes the posterior pdf of the states conditioned on the measurements y_k. A similar form of the Bayes' rule can be used to compute the posterior pdf $\mathsf{P}(\theta_i|y_k),\ i = 1, 2, \ldots, p$ of the uncertain parameters.

There are two main challenges involved in solving the Bayes' rule in (2) for the uncertain nonlinear system (1). The first challenge concerns efficient uncertainty propagation through the nonlinear system dynamics. The gPC framework [6] is used for uncertainty propagation (see Section III). The second challenge arises from computing the prior, likelihood, and evidence in (2). Analytical expressions cannot be obtained for the pdfs due to the nonlinearity of the system dynamics. The HF algorithm is used to approximate these distributions (see Section IV).

III. Uncertainty propagation using Polynomial Chaos

The generalized polynomial chaos (gPC) framework [6] is used for efficient propagation of the probabilistic uncertainties in θ and x_0. In the gPC framework, a stochastic variable $\psi(\xi)$ is expressed as an infinite series expansion of orthogonal polynomial basis functions

$$\psi(\xi) \coloneqq \sum_{j=0}^{\infty} a_j \varphi_j(\xi),$$

where a_j are the expansion coefficients and $\varphi_j(\xi), j \in \mathbb{N}_0$ are the basis functions with maximum degree m with respect to the standard random variables ξ. The basis functions belong to the Wiener-Askey scheme of polynomials [6], which consist of orthogonal basis functions in $L^2\left\{\Omega, \mathcal{F}, \mathsf{P}\right\}$ defined on the support space of the random variables ξ. The orthogonality of the basis functions implies that $\left\langle \varphi_i\left(\xi\right), \varphi_j\left(\xi\right)\right\rangle = \left\langle \varphi_i^2\left(\xi\right)\right\rangle \delta_{ij}$, where $\left\langle g_1\left(\xi\right), g_2\left(\xi\right)\right\rangle = \int_\Omega g_1\left(\xi\right) g_2\left(\xi\right) \mathsf{P}\left(\xi\right) d\xi$ is the inner product induced by $\mathsf{P}\left(\xi\right)$ and δ_{ij} is the Kronecker delta function. For computational tractability, the expansion is truncated after $l+1$ terms, where $l+1 = \frac{(n_\xi+m)!}{n_\xi! m!}$. Thus, $\psi\left(\xi\right)$ is approximated as

$$\psi\left(\xi\right) \approx \hat{\psi}\left(\xi\right) = \sum_{j=0}^{l} a_j \varphi_j\left(\xi\right) = a\Lambda^\top\left(\xi\right), \qquad (3)$$

where $a := \left[a_0, \ldots, a_l\right]$ and $\Lambda\left(\xi\right) := \left[\varphi_0\left(\xi\right), \ldots, \varphi_l\left(\xi\right)\right]^\top$.

The expansion coefficients are defined by

$$a_j = \frac{\left\langle \hat{\psi}\left(\xi\right), \varphi_j\left(\xi\right)\right\rangle}{\left\langle \varphi_j\left(\xi\right), \varphi_j\left(\xi\right)\right\rangle}, \quad \forall\, j = 0, \ldots, l.$$

The inner products in the above equation can be computed by evaluating the integrals analytically (the Galerkin projection) when the system is of polynomial form [19], or through sample-based methods such as the probabilistic collocation methods (PCM) (e.g., see [20]). A brief description of both methods is provided below (see [21] for further details).

A. Galerkin Projection

The ith state in (1) can be approximated using the polynomial chaos expansion (3). The system equation for each approximated uncertain state $\hat{x}_{i,k}\left(\xi\right)$ takes the form

$$\sum_{j=0}^{l} \tilde{x}_{i,j,k} \varphi_j\left(\xi\right) = f_i\left(\tilde{x}_{1,k-1}\Lambda^\top\left(\xi\right), \ldots, \tilde{\theta}_1\Lambda^\top\left(\xi\right), \ldots, u\right), \quad \forall\, i = 1, \ldots, n, \qquad (4)$$

where $\tilde{x}_{i,k}$ and $\tilde{\theta}_i$ denote the vector of expansion coefficients. Since the system equations (1) are expressed as polynomial functions in states, the Galerkin projection method is applied to (4) to obtain a set of discrete-time, deterministic equations for describing the dynamics of the coefficients of the PC expansion for each state $\hat{x}_{i,k}$. The Galerkin projection is performed by computing the following inner product [19]

$$\int_\Omega f_i\left(\tilde{x}_{1,k-1}\Lambda^\top\left(\xi\right), \ldots, \tilde{\theta}_1\Lambda^\top\left(\xi\right), \ldots, u\right) \mathsf{P}\left(\xi\right).$$

The above projection ensures that the approximation error between the true states and their respective polynomial chaos approximations is orthogonal to the functional space spanned by $\varphi_j\left(\xi\right)$ [21]. The orthogonality of the basis functions results in the following set of discrete-time equations that describes the dynamics of the PC expansion coefficients for each approximated state $\hat{x}_{i,k}$

$$\tilde{x}_{i,k} = \tilde{f}_i\left(\tilde{x}_{1,k-1}, \ldots, \tilde{\theta}_1, \ldots, u\right), \quad \forall\, i = 1, \ldots, n, \qquad (5)$$

where $\tilde{f}_i\left(\cdot\right)$ describes the dynamics of the expansion coefficients of the ith state.

B. Probabilistic Collocation Method

When the system equations are nonpolynomial in states, it is impractical to use the Galerkin method.[1] In this case, the PCM is used for determining the expansion coefficients by requiring the PC expansions $\hat{\psi}\left(\xi\right)$ be exact at some chosen collocation points. To this end, the approximation residual is defined as

$$\varrho\left(a, \xi\right) := \psi - \hat{\psi}\left(\xi\right) = \psi - \sum_{j=0}^{l} a_j \varphi_j\left(\xi\right).$$

The coefficients $\left\{a_j\right\}_{j=0}^{l}$ are computed such that $\varrho\left(a, \xi\right)$ is orthogonal to $\varphi_j\left(\xi\right)$ [20]

$$\int \varrho\left(a, \xi\right) \varphi_j\left(\xi\right) \mathsf{P}\left(\xi\right) d\xi = 0, \quad \forall\, j = 0, \cdots, l.$$

The above integral can be approximated using a quadrature method

$$\sum_{o=0}^{n_c} \omega^{(o)} \varrho(a, \xi^{(o)}) \varphi_j(\xi^{(o)}) \mathsf{P}(\xi^{(o)}) = 0, \ \forall\, j = 0, \ldots, l, \qquad (6)$$

where $\omega^{(o)}$ denotes the weights of the quadrature approximation; $n_c \in \mathbb{N}$ is the number of collocation points; and $\xi^{(o)}$ denotes the samples of the standard random variable ξ drawn from the pdf $\mathsf{P}\left(\xi\right)$. Using (6), the PC expansion coefficients a can be estimated by computing $\varrho\left(a, \xi^{(o)}\right)$ at n_c collocation points. The collocation points can be chosen either deterministically as the roots of the polynomial basis function of order $\left(n_\xi + 1\right)$, or by random sampling of the known distributions of uncertainties.

IV. GPC-Based Histogram Filter

The proposed gPC-based histogram filter, which approximates the posterior pdf of the stochastic states and uncertain parameters, is presented in this section. The HF algorithm discretizes the support of the pdf of the states and parameters to obtain a closed-form approximation of the Bayes' rule [5].

A. State Estimation

At time k, the prior pdf of states is obtained as follows. First, the coefficients of the PC expansions of states, i.e., $\left\{\tilde{x}_{i,k|k-1}\right\}_{i=1}^{n}$, are computed by solving (5). Subsequently, n_p realizations of the approximated stochastic states $\left\{\hat{x}_{i,k|k-1}\right\}_{i=1}^{n}$ are obtained by drawing n_p random samples from $\mathsf{P}\left(\xi\right)$ and $\mathsf{P}\left(w\right)$

$$\begin{aligned} &\xi^{(j)} \sim \mathsf{P}\left(\xi\right), \ j = 1, \ldots, n_p, \\ &w_{k-1}^{(j)} \sim \mathsf{P}\left(w\right), \ j = 1, \ldots, n_p, \\ &\hat{x}_{i,k|k-1}^{(j)} = \tilde{x}_{i,k|k-1}\Lambda^\top(\xi^{(j)}) + w_{k-1}^{(j)}, \quad \forall\, i = 1, \ldots, n. \end{aligned}$$

A multivariate histogram of the stochastic states $\{\hat{x}_{i,k|k-1}^{(j)}\}_{j=1}^{n_p}$ can now be constructed using n_b bins. Let $\left\{N_{c,i}\right\}_{i=1}^{n_b}$ represent the number of observations of the states in each bin of the histogram, with c_i being the center of the

[1]When transforming a general nonlinear model to its polynomial-in-states form, the resulting equations may not be polynomial in the artificial states.

ith bin. The number of bins is chosen as $n_b = \lceil \sqrt{n_p} \rceil$, where $\lceil \cdot \rceil$ represents the ceiling operator.[2] Given $\{N_{c,i}, c_i\}_{i=1}^{n_b}$, the prior pdf with respect to the bin centers c_i is computed as

$$\mathsf{P}\left(x_k = c_i | y_{k-1}\right) = \frac{N_{c,i}}{\sum_{i=1}^{N_b} N_{c,i} b_i}, \ \forall \ i = 1, \dots, n_b, \quad (7)$$

where b_i denotes the interval of each bin. On the other hand, the likelihood of y_k with respect to c_i is defined by[3]

$$\begin{aligned} v_{i,k} &= y_k - h\left(c_i\right), \\ \mathsf{P}\left(y_k | x_k = c_i\right) &= \mathcal{N}\left(v_{i,k}; 0, R\right), \quad \forall \ i = 1, \dots, n_b. \end{aligned} \quad (8)$$

Now, the expressions (7) and (8) can be used to estimate the posterior pdf of the states as

$$\mathsf{P}\left(x_k = c_i | y_k\right) = \frac{\mathsf{P}\left(x_k = c_i | y_{k-1}\right) \mathsf{P}\left(y_k | x_k = c_i\right)}{\mathsf{P}\left(y_k\right)} \quad (9)$$

with

$$\mathsf{P}\left(y_k\right) = \sum_{i=1}^{n_b} b_i \mathsf{P}\left(x_k = c_i | y_{k-1}\right) \mathsf{P}\left(y_k | x = c_i\right).$$

Note that the area under the pdf $\mathsf{P}\left(x_k | y_k\right)$ is unity.

Since the posterior pdf of states evolves in time (as described by (9)), the coefficients of the PC expansions of states must be updated accordingly. An optimization problem is formulated to recompute the coefficients of the PC expansions based on the histogram of the posterior pdf of states. Using the gPC framework, the qth moment of $\hat{x}_{i,k|k}$ in terms of its PC expansion coefficients is given by

$$v_q\left(\hat{x}_{i,k}\right) = \sum_{j_1=0}^{l} \cdots \sum_{j_q=0}^{l} \tilde{x}_{i,j_1,k} \cdots \tilde{x}_{i,j_q,k} \left\langle \varphi_{j_1}\left(\xi\right) \cdots \varphi_{j_q}\left(\xi\right) \right\rangle. \quad (10)$$

Using the histogram of states, the qth posterior moment of $\hat{x}_{i,k}$ is approximated as

$$\mathcal{M}_{x,q}\left(\hat{x}_{i,k}\right) := \mathbf{E}^q\left[\hat{x}_{i,k}\right] \approx \sum_{j=1}^{n_p} \omega_x^{(j)} \left(\hat{x}_{i,k}^{(j)}\right)^q, \quad (11)$$

where $\omega_x^{(j)}$ is the normalized posterior pdf of the jth sample of the states.

The coefficients of the PC expansion of the ith state is computed by minimizing the sum of squared error between the moments obtained using (10) and (11)

$$\tilde{x}_{i,k|k}^* := \arg \min_{\tilde{x}_{i,k}} \sum_{q=1}^{m+1} \parallel v_q\left(\hat{x}_{i,k}\right) - \mathcal{M}_{x,q}\left(\hat{x}_{i,k}\right) \parallel^2. \quad (12)$$

The updated coefficients obtained from (12) are used as the initial condition in (5) for the next time step.

[2]This relation gives the minimum number of required bins. A larger number of bins can be used to enhance the pdf approximations.

[3]The measurement error is assumed to have a zero-mean Normal distribution with variance R. However, the proposed estimation algorithm is valid for arbitrary type measurement noise distributions.

B. Parameter Estimation

The Bayes' rule is applied to estimate the uncertain parameters. Due to the time-invariant nature of parametric uncertainties in (1), the following holds

$$\mathsf{P}\left(\theta_{i,k} | y_{k-1}\right) = \mathsf{P}\left(\theta_{i,k-1} | y_{k-1}\right), \ \forall \ i = 1, \dots, p.$$

The samples of the parameters are available from the previous sampling instant. Thus, for $\{\theta_i\}_{i=1}^{p}$ at time k

$$\hat{\theta}_{i,k|k-1}^{(j)} = \hat{\theta}_{i,k-1|k-1}^{(j)}, \ \forall \ j = 1, \dots, n_p$$

such that

$$\mathsf{P}\left(\theta_{i,k} = \hat{\theta}_i^{(j)} | y_{k-1}\right) = \mathsf{P}\left(\theta_{i,k-1} = \hat{\theta}_i^{(j)} | y_{k-1}\right), \quad (13)$$

where n_p denotes the number of random realizations of the uncertain parameters.

The likelihood of y_k for every $\hat{\theta}_i^{(j)}$ is given by

$$\begin{aligned} v_k^{(j)} &= y_k - \hat{y}_k(\theta_i = \hat{\theta}_i^{(j)}), \\ \mathsf{P}\left(y_k | \hat{\theta}_i^{(j)}\right) &= \mathcal{N}\left(v_k^{(j)}; 0, R\right). \end{aligned} \quad (14)$$

Thus, the posterior pdf of the parameters can be computed by using (13) and (14) as

$$\mathsf{P}\left(\theta_i = \hat{\theta}_i^{(j)} | y_k\right) = \frac{\mathsf{P}\left(y_k | \hat{\theta}_i^{(j)}\right) \mathsf{P}\left(\theta_i = \hat{\theta}_i^{(j)} | y_{k-1}\right)}{\mathsf{P}\left(y_k\right)},$$

where

$$\begin{aligned} \mathsf{P}\left(y_k\right) &= \sum_{j=2}^{n_p} \delta\theta_i^{(j)} \mathsf{P}\left(y_k | \hat{\theta}_i^{(j)}\right) \mathsf{P}\left(\theta_i = \hat{\theta}_i^{(j)} | y_{k-1}\right), \\ \delta\theta_i^{(j)} &= \hat{\theta}_i^{(j)} - \hat{\theta}_i^{(j-1)}. \end{aligned}$$

Similar to (12), the coefficients of PC expansions of the parameters must be updated by minimizing the sum of squared error between the moments obtained using the gPC framework and the moments obtained using the histogram of the posterior pdf of the parameters

$$\tilde{\theta}_{i,k|k}^* := \arg \min_{\tilde{\theta}_{i,k}} \sum_{q=1}^{m+1} \parallel v_q(\theta_i) - \mathcal{M}_q(\theta_i) \parallel^2, \quad (15)$$

where

$$v_q(\theta_i) = \sum_{j_1=0}^{l} \cdots \sum_{j_q=0}^{l} \tilde{\theta}_{i,j_1} \cdots \tilde{\theta}_{i,j_q} \left\langle \varphi_{j_1}(\xi) \cdots \varphi_{j_q}(\xi) \right\rangle,$$

$$\mathcal{M}_q(\theta_i) := \mathbf{E}^q\left[\theta_i\right] \approx \sum_{j=1}^{n_p} \omega_\theta^{(j)} \left(\theta_i^{(j)}\right)^q,$$

and $\omega_{\theta_i}^{(j)}$ is the normalized posterior pdf of the jth sample of θ_i.

A sufficiently large number of PC terms may be needed to obtain good estimates for the arbitrary type pdfs of the stochastic states and uncertain parameters.

V. Case Study: A Continuous Bioreactor

The performance of the gPC-based histogram filter is demonstrated on a benchmark continuous bioreactor [17]. The continuous-time system dynamics are described by

$$dh = \Big(\frac{F_{in}}{\pi r^2} - \frac{k_1}{\pi r^2}\sqrt{h}\Big)dt + \sigma_h dw_h(t) \tag{16a}$$

$$dX = \Big(-\frac{F_{in}}{\pi r^2 h}X + \mu X\Big)dt + \sigma_X dw_X(t) \tag{16b}$$

$$dS = \Big(\frac{F_{in}}{\pi r^2 h}(S_f - S) - \frac{1}{Y_{X|S}}\mu X\Big)dt + \sigma_S dw_S(t) \tag{16c}$$

$$dP = \Big(-\frac{F_{in}}{\pi r^2 h}P + (\alpha\mu + \beta)X\Big)dt + \sigma_P dw_P(t), \tag{16d}$$

where h is the level of the reactor; and X, S, and P are the biomass concentration, substrate concentration, and product concentration in the bioreactor. The terms $w_h(t)$, $w_X(t)$, $w_S(t)$, and $w_P(t)$ are independent Wiener processes with variances $\sigma_i,\ i = \{h, X, S, P\}$ as given in Q in Table I. The inlet flow rate F_{in} and the inlet substrate concentration S_f are the inputs, and r is the radius of the reactor tank. $Y_{X|S}$ is the cell-biomass yield, and α and β are the yield parameters for P. μ denotes the specific growth rate with substrate inhibition

$$\mu = \frac{\mu_m S}{K_m + S},$$

where μ_m is the maximum specific growth rate. The system parameters can be found in [17].

The bioreactor is run for a period of 12 hr with the steady-state operating point chosen as the initial conditions. The states h, S, and P are measured at regular sampling time intervals of 0.25 hr. The measurements are corrupted by white noise $v \sim \mathcal{N}(0, R)$. The parameter μ_m is assumed to be uncertain with distribution $\mu_m \sim \mathcal{N}(\mu_{m,0}, \sigma_m)$. The initial states are described by Normal distributions $\mathcal{N}(m_{i,0}, \sigma_{i,0}),\ i = \{h, X, S, P\}$, where $m_{i,0}$ and $\sigma_{i,0}$ are, respectively, the mean and variance of the initial states. The properties of the system uncertainties and disturbances are listed in Table I.

As the time-invariant parametric uncertainty and uncertain initial states are all described by Normal distributions, Hermite polynomials are chosen as the basis functions in the PC expansions. The order of the polynomial basis functions is selected as $m = 3$. There are five uncertain variables (i.e., $n_\xi = 5$); hence, the total number of terms in the PC expansions is $l = 56$ (see Section III). The model equations (16) consist of nonpolynomial expressions, which are transformed to a polynomial-in-states form through state-lifting [18]. This results in three extra artificial states, leading to a total of seven dynamic system equations (not shown here). A combination of the Galerkin projection and the probabilistic collocation method is used to determine the coefficients of the PC expansions.

TABLE I: Properties of system uncertainties and disturbances

Variable	Covariance/variance
Q	$10^{-3} \times \text{diag}\,[11.1\ \ 0.7\ \ 139.2\ \ 0.12]$
R	$10^{-3} \times \text{diag}\,[0\ \ 5.6\ \ 15.6\ \ 0.1]$
$\{\mu_{m,0}, \sigma_m\}$	$\{0.48, 1 \times 10^{-4}\}$
$\{m_{X,0}, \sigma_{X,0}\}$	$\{7.04, 0.12\}$
$\{m_{S,0}, \sigma_{S,0}\}$	$\{2.40, 0.015\}$
$\{m_{P,0}, \sigma_{P,0}\}$	$\{24.87, 1\}$
$\{m_{h,0}, \sigma_{h,0}\}$	$\{1, 0.003\}$

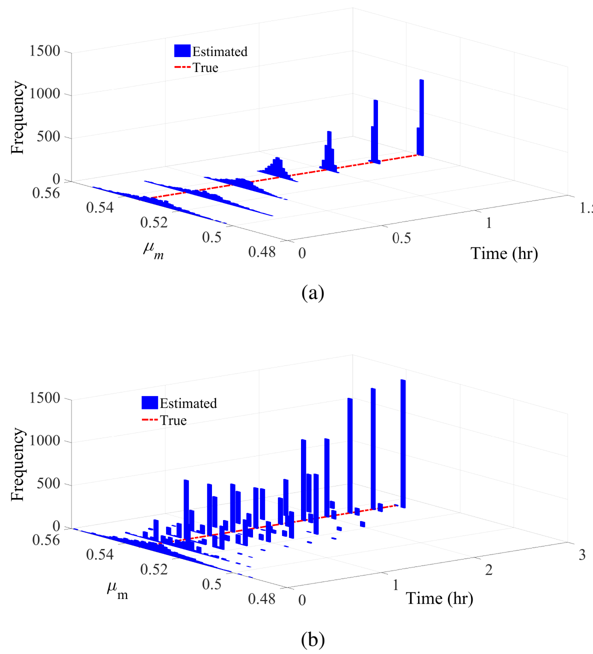

Fig. 1: Evolution of the posterior pdf of the uncertain parameter μ_m estimated using (a) the gPC-based histogram filter and (b) the SIR particle filter.

Fig. 1a shows the evolution of the estimated posterior pdf of μ_m. The estimated parameter values converge to the true value of μ_m, and the pdf of the uncertain parameter reduces to a Dirac-delta function within a few time steps. The performance of the gPC-based histogram filter is compared to that of a SIR particle filter [2] with 1000 MC particles. The process conditions are identical to those used in the gPC-based histogram filter. The evolution of the pdf of μ_m estimated by the SIR particle filter, as shown in Fig. 1b, is similar to that estimated by the gPC-based histogram filter. However, the parameter estimates of the SIR particle filter are slightly biased with respect to the true parameter value. Further, the gPC-based histogram filter exhibits a faster convergence to the true parameter value.

Fig. 2 shows the true profile of P as well as the mean of its posterior pdfs estimated by the gPC-based histogram filter and the SIR particle filter. The gPC-based histogram filter provides reasonably accurate estimates for the state P. Table II lists the root-mean-squared errors (RMSEs) of the state estimates obtained by the gPC-based histogram filter and the SIR particle filter based on 100 Monte Carlo simulations of both estimators. The RMSE results indicate

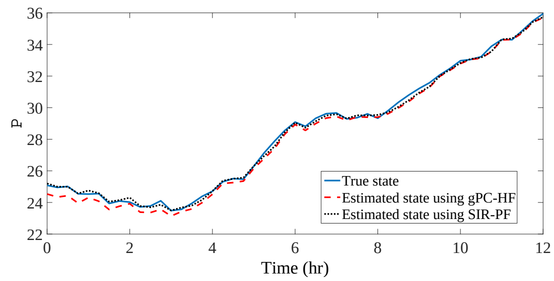

Fig. 2: Comparison of the true state P and the mean of the posterior pdf of P estimated by the gPC-based histogram filter and the SIR particle filter.

that the gPC-based histogram filter slightly outperforms the SIR particle filter, except for the case of estimating the state P. This is due to the slower convergence rate of the gPC-based histogram filter for the estimation of P (see Fig. 2).

The performance of the two nonlinear estimation algorithms is also compared in terms of their computational requirements. On a desktop with a 3.6 GHz Intel Core-i7 processor and 8GB RAM, the CPU time for the prediction step of the gPC-based histogram filter is 0.051 s, while that of the SIR particle filter is 1.32 s. This suggests that the gPC framework is a computationally efficient approach for propagation of the probabilistic uncertainties. On the other hand, the update step in the gPC-based histogram filter is computationally more expensive (0.31 s) than that of the SIR particle filter (0.05 s). This is due to the need to solve a nonlinear least squares problem (i.e., (12) and (15)). The computational advantage of the SIR particle filter in the update step results from the fact that it uses the prior pdfs as the importance distribution to sample the particles of the posterior pdf. Overall, the gPC-based histogram filter is approximately five times faster than the SIR particle filter.

VI. Conclusions

This paper presents a polynomial chaos-based histogram filter for constructing the posterior pdf of stochastic states and uncertain parameters. The advantages of the proposed nonlinear estimation algorithm include accounting for time-invariant system uncertainties and time-varying system disturbances, utilizing the computationally efficient framework of polynomial chaos for propagating the probabilistic uncertainties, and estimating the full posterior pdf of states and parameters. In future, the conditions for unbiased estimates and convergence properties of the polynomial-chaos based histogram filter will be investigated.

TABLE II: RMSE values of state estimates obtained by the gPC-based histogram filter and SIR particle filter

State	gPC-based Histogram Filter	SIR Particle Filter
X	0.290	0.328
S	0.057	0.059
P	0.199	0.124
h	0.013	0.016

References

[1] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, “Novel approach to nonlinear/non-Gaussian Bayesian state estimation,” *IEE Proceedings on Radar and Signal Processing*, vol. 140, no. 2, pp. 107–113, 1993.

[2] M. S. Arulampalam, S. Maskell, and T. Clapp, “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking,” *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.

[3] S. C. Patwardhan, S. Narasimhan, J. Prakash, R. B. Gopaluni, and S. L. Shah, “Nonlinear Bayesian state estimation: A review of recent developments,” *Control Engineering Practice*, vol. 20, pp. 933–953, 2012.

[4] G. Evensen, *Data Assimilation– The ensemble Kalman filter*. Springer-Verlag, Heidelberg, 2007.

[5] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. The MIT Press, Cambridge, 2005.

[6] D. Xiu and G. E. Karniadakis, “The Wiener–Askey polynomial chaos for stochastic differential equations,” *SIAM Journal on Computing*, vol. 24, no. 2, pp. 619–644, 2002.

[7] J. Fisher and R. Bhattacharya, “Linear quadratic regulation of systems with stochastic parameter uncertainties,” *Automatica*, vol. 45, no. 12, pp. 2831–2841, 2009.

[8] A. Mesbah, S. Streif, R. Findeisen, and R. D. Braatz, “Stochastic nonlinear model predictive control with probabilistic constraints,” in *Proceedings of the American Control Conference*, 2014, pp. 2413 – 2419.

[9] ——, “Active fault diagnosis for nonlinear systems with probabilistic uncertainties,” in *Proceedings of 19th IFAC World Congress*, 2014.

[10] A. Mesbah and S. Streif, “A probabilistic approach to robust optimal experiment design with chance constraints,” in *Proceedings of the 9th International Symposium on Advanced Control of Chemical Processes (ADCHEM)*, 2015, pp. 100–105.

[11] J. Li and D. Xiu, “A generalized polynomial chaos based ensemble Kalman filter with high accuracy,” *Journal of Computational Physics*, vol. 228, no. 15, pp. 5454–5469, 2009.

[12] E. D. Blanchard, A. Sandu, and C. Sandu, “A polynomial chaos-based Kalman filter approach for parameter estimation of mechanical systems,” *Journal of Dynamic Systems, Measurement and Control*, vol. 132, no. 6, pp. 061 404–1–061 404–18, 2010.

[13] P. Dutta and R. Bhattacharya, “Nonlinear estimation of hypersonic state trajectories in Bayesian framework with polynomial chaos,” *Journal of Guidance, Control, and Dynamics*, vol. 33, no. 6, pp. 1765–1778, 2010.

[14] U. Konda, P. Singla, T. Singh, and P. Scott, “State uncertainty propagation in the presence of parametric uncertainty and additive white noise,” *Journal of Dynamic Systems, Measurement and Control*, vol. 133, no. 5, pp. 051 009–1–10, 2011.

[15] R. Madankan, P. Singla, T. Singh, and P. D. Scott, “Polynomial chaos-based Bayesian approach for state and parameter estimations,” *Journal of Guidance Control and Dynamics*, vol. 36, no. 4, pp. 1058–1074, 2013.

[16] R. Madankan, S. Pouget, P. Singla, M. Bursik, J. Dehn, M. Jones, A. Patra, M. Pavalonis, E. B. Pitman, T. Singh, and P. Webley, “Computation of probabilistic hazard maps and source parameter estimation for volcanic ash transport and dispersion,” *Journal of Computational Physics*, vol. 271, pp. 39 – 59, 2014.

[17] P. Agrawal, G. Koshy, and M. Ramseier, “Algorithm for operating a fed-batch fermentor at optimum specific-growth rate,” *Biotechnology and Bioengineering*, vol. 33, no. 1, pp. 115–125, 1989.

[18] A. Papachristodoulou and S. Prajna, “Analysis of non-polynomial systems using the sum of squares decomposition,” in *Positive Polynomials in Control*, ser. Lecture Notes in Control and Information Science, D. Henrion and A. Garulli, Eds. Springer, Berlin Heidelberg, 2005, pp. 23–43.

[19] R. Ghanem and P. Spanos, *Stochastic Finite Elements– A Spectral Approach*. Springer-Verlag, New York, 1991.

[20] M. A. Tatang, W. Pan, R. G. Prinn, and G. J. McRae, “An efficient method for parametric uncertainty analysis of numerical geophysical models,” *Journal of Geophysical Research: Atmospheres*, vol. 102, no. D18, pp. 21 925–21 932, 1997.

[21] K.-K. K. Kim, D. E. Shen, Z. K. Nagy, and R. D. Braatz, “Wiener’s polynomial chaos for the analysis and control of nonlinear dynamical systems with probabilistic uncertainties,” *IEEE Control Systems Magazine*, vol. 33, no. 5, pp. 58–67, 2013.