18th International Conference on Information Fusion
Washington, DC - July 6-9, 2015

Nonlinear Bayesian Filtering Based on Fokker-Planck Equation and Tensor Decomposition

Yifei Sun and Mrinal Kumar

***Abstract*— A nonlinear Bayesian filter is proposed in this paper for a general nonlinear system of continuous time dynamics and discrete time measurements. In this filter, a transient Fokker-Planck equation solver based on tensor decomposition is used for propagating the conditional state probability density function (PDF) in conjunction with a measurement update via Bayes' rule. This filter is not restricted by assumptions of linearity or Gaussianity since it relies on the exact state PDF which captures the entire information of the underlying uncertainty. Moreover, it is suitable for system of high-dimensional state space by virtue of the efficient tensor decomposition scheme, which enables the computational efforts for solving the state PDF grow benignly with dimensionality. This is possible because every dimension of the state space as well as the time domain is separated from each other in the solution process, as a result of which originally expensive high-dimensional operations are decoupled into a series of simple one-dimensional operations. Numerical examples are provided to demonstrate the advantages of the proposed filter over the extended Kalman filter for state estimation.**

I. Introduction

The knowledge of the state is of great value for a dynamical system in that it is essential for monitoring and control. However, in reality the true system model is seldom known exactly and subject to certain noise perturbation, the initial condition is only given stochastically as a probability density function (PDF) rather than deterministically, and the limited measurements are corrupted by noise as well. Consequently, a filter is required for state estimation [1], [2]. For linear dynamical system with linear measurement model and Gaussian perturbations, the Kalman filter [3] represents the optimal estimator [4]. For general nonlinear systems, the design of optimal nonlinear filter is difficult since it usually entails infinite number of parameters if one resorts to methods based on moment evolution. As a compromise, several finite dimensional suboptimal filters have been developed [5], among which the extended Kalman filter (EKF) is arguably the most popular. The essential idea of EKF is to use the Jacobian matrices of the nonlinear dynamics and measurement model for error dynamics and gain calculation. Although not optimal, the EKF has been successfully applied to various nonlinear systems over the past several decades. Note however that it can perform poorly for highly nonlinear cases with large uncertainty in initial condition and long propagation time [6].

The recursive Bayesian approach represents a general way to describe the optimal nonlinear filtering problem, where the key is to propagate the conditional state probability density function (PDF) governed by the Fokker-Planck equation (FPE) and update the PDF using the measurement data by the Bayes' rule. Filters based on this idea have been developed, for example, in Refs. [7], [8], [9], which are not restricted by assumptions of linearity or Gaussianity since they rely on the exact state PDF which captures the entire information of the underlying uncertainty. On the other hand, the high accuracy is achieved at the cost of usually high computational efforts since the FPE, a second order parabolic partial differential equation (PDE) need to be solved for each prediction step. This problem becomes even more severe when the dimensionality of the underlying state space is large [10], [11], since solving a PDE like FPE in general suffers from the well-known curse of dimensionality [12], i.e., the degrees of freedom (DOF) of the approximation, or number of unknowns grow exponentially with respect to the dimensionality [13], [14], [15]. Numerical methods such as the meshless partition of unity finite element method (PUFEM) has been employed to counter this problem with moderate success [16], [17], [18]. In this paper, the transient FPE solver developed in Refs. [19], [20] will be adopted in the prediction step of the nonlinear Bayesian filter. To tackle the dimensionality issue for solving FPEs, this solver combines the Chebyshev spectral method and a tensor decomposition approach to drastically reduce the DOF required for maintaining accuracy of the solution. This is due to the spectral accuracy of the Chebyshev differentiation and the fact that every dimension of the state space as well as the time domain is separated from each other in the solution process, as a result of which originally expensive high-dimensional operations are decoupled into a series of simple one-dimensional operations. This transient solver has been successfully applied to solving high dimensional FPEs (up to 14D transient problem) encountered in various areas including nonlinear vibrations [21], [22], polymeric fluids [23] and orbital mechanics [24].

The remainder of this paper is organized as follows: section II states the problem of nonlinear filtering, followed by the description of extended Kalman filter in section III. The proposed nonlinear Bayesian filter's scheme and numerical examples are given in section IV and V respectively. Finally, a summary and future research directions are provided in

This work was supported by the National Science Foundation under Grant No. ECCS-1254244 and the NASA Marshall Space Flight Center Grant No. NNM13AA07G

Y. Sun is a post-doctoral associate in the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32608, USA yfsun@ufl.edu

M. Kumar is an assistant professor in the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32608, USA mrinalkumar@ufl.edu

section VI.

II. Problem Statement

Consider the following continuous nonlinear dynamical system with discrete time measurements:

$$d\mathbf{x} = \mathbf{f}(t, \mathbf{x})dt + \mathbf{g}(t)d\mathbf{B}(t), \quad \mathbf{x} \in \mathbb{R}^P, \tag{1}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k, \tag{2}$$

where $B(t)$ denotes the Brownian motion process with zero mean and covariance $\mathbf{Q}t$. $\mathbf{f}(t, \mathbf{x}), \mathbf{g}(t)$ are deterministic functions, and $\mathbf{y}_k$ is the k^{th} measurement vector of size $Q \times 1$. Let $\mathbf{v}_k$ be a zero mean Gaussian white noise, such that

$$E\{\mathbf{v}_k \mathbf{v}_j^T\} = \begin{cases} 0 & k \neq j, \\ \mathbf{R} & k = j. \end{cases} \tag{3}$$

Further, it is assumed that $\mathbf{v}_k$ is uncorrelated with $d\mathbf{B}(t)$. Note that the $\mathbf{Q}$ of process noise and measurement noise covariance $\mathbf{R}$ could change with time in practice, but here we assume they are constant for simplicity. The time varying state PDF $\mathcal{W}(t, \mathbf{x})$ of Eq.1 is governed by the Fokker-Planck equation:

$$\frac{\partial}{\partial t}\mathcal{W}(t, \mathbf{x}) = \mathcal{L}_{\mathcal{FP}}[\mathcal{W}(t, \mathbf{x})], \tag{4}$$

where $\mathcal{L}_{\mathcal{FP}}$ is the Fokker-Planck operator given by

$$\mathcal{L}_{\mathcal{FP}} = \left[-\sum_{i=1}^{\mathrm{P}} \frac{\partial}{\partial x_i} \mathrm{D}_i^{(1)}(\cdot) + \sum_{i=1}^{\mathrm{P}} \sum_{j=1}^{\mathrm{P}} \frac{\partial^2}{\partial x_i \partial x_j} \mathrm{D}_{ij}^{(2)}(\cdot) \right], \tag{5}$$

$$\mathrm{D}^{(1)}(t, \mathbf{x}) = \mathbf{f}(t, \mathbf{x}), \quad \mathrm{D}^{(2)}(t, \mathbf{x}) = \frac{1}{2}\mathbf{g}(t, \mathbf{x}) Q \mathbf{g}^{\mathrm{T}}(t, \mathbf{x}).$$

The filtering process is to estimate the state $\mathbf{x}$, given the general knowledge of the dynamics (since in reality the $\mathbf{f}(t, \mathbf{x}), \mathbf{g}(t)$ in Eq.1 are usually not known exactly, and $\mathbf{Q}$ serves as a tuning parameter accordingly) and the initial condition (whose uncertainty is quantified as $\mathcal{W}(t_0, \mathbf{x}) = \mathcal{W}_0(\mathbf{x})$), combined with the measurement model of Eq.2, $\mathbf{y}_k$ and $\mathbf{R}$.

III. Extended Kalman Filter

As an extension of the Kalman filter for linear system, for generally nonlinear system the extended Kalman filter was developed, which uses the Jacobian matrices of the nonlinear dynamics and measurement model for error dynamics and gain calculation. This assumes the $\mathbf{f}$ and $\mathbf{h}$ in Eqs.1 and 2 are continuously differentiable and the true state is sufficiently close to the estimated state [6]. With these adjustments, the EKF follows the same structure as the Kalman filter.

Let $\hat{\mathbf{x}}$ denote the estimate of the state $\mathbf{x}$, then define the estimate error as $\tilde{\mathbf{x}} = \hat{\mathbf{x}} - \mathbf{x}$, and use $+, -$ in the superscript to describe the prior and posterior versions of a quantity. Define the following error covariances:

$$\mathbf{P}_k^- = E\{\tilde{\mathbf{x}}_k^- \tilde{\mathbf{x}}_k^{-T}\}, \;\; \mathbf{P}_k^+ = E\{\tilde{\mathbf{x}}_k^+ \tilde{\mathbf{x}}_k^{+T}\}. \tag{6}$$

The essential steps of EKF is described as follows:

- *Initialization:*

$$\hat{\mathbf{x}}(t_0) = \hat{\mathbf{x}}_0, \;\; \mathbf{P}_0 = E\{\tilde{\mathbf{x}}(t_0)\tilde{\mathbf{x}}^T(t_0)\}, \tag{7}$$

where $\hat{\mathbf{x}}(t_0)$ and $\mathbf{P}_0$ capture the mean and covariance of the PDF for initial condition.

- *Propagation:*

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{f}(t, \hat{\mathbf{x}}(t)), \tag{8}$$

$$\dot{\mathbf{P}}(t) = \mathbf{F}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}^T(t) + \mathbf{g}(t)\mathbf{P}(t)\mathbf{Q}\mathbf{g}^T(t), \tag{9}$$

$$\mathbf{F}(t) = \left. \frac{\partial \mathbf{f}(t, \mathbf{x})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}(t)}. \tag{10}$$

- *Update:*

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + \mathbf{K}_k[\mathbf{y}_k - \mathbf{h}(\hat{\mathbf{x}}_k^-)], \tag{11}$$

$$\mathbf{P}_k^+ = [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k(\hat{\mathbf{x}}_k^-)]\mathbf{P}_k^-, \tag{12}$$

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T(\hat{\mathbf{x}}_k^-)[\mathbf{H}_k(\hat{\mathbf{x}}_k^-)\mathbf{P}_k^- \mathbf{H}_k^T(\hat{\mathbf{x}}_k^-) + \mathbf{R}]^{-1}, \tag{13}$$

$$\mathbf{H}_k(\hat{\mathbf{x}}_k^-) = \left. \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_k^-}. \tag{14}$$

After the initialization, the state estimate $\hat{\mathbf{x}}$ and error covariance $\mathbf{P}$ are propagated using Eqs.8 to 10 until the first measurement arrives. Then the update of Eqs.11 to 14 takes place. The above process repeats sequentially until all measurement information is used. If a measurement is available at the beginning, the update step is immediately taken.

IV. A FPE Based Nonlinear Bayesian Filter

In EKF it is assumed that the conditional state PDF keeps Gaussian even for nonlinear dynamics [2], which may lead to large error in implementation. To estimate the real underlying evolution of conditional state PDF, the FPE based Bayesian filter is developed, in which the FPE is solved in the propagation step to obtain the prior prediction. The measurement data is then incorporated using the Bayes' rule in the update step to compute the posterior. This becomes the initial condition for the next propagation step. These propagation and update steps repeat sequentially until all measurement data are used.

Define $\mathcal{Y}^k = \{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_k\}$ which contains all the measurement information up to the k^{th} measurement. Then the propagation step of the Bayesian filter is given by the Chapman-Kolmogorov equation (CKE):

$$\begin{aligned} \mathcal{W}(\hat{\mathbf{x}}_k^-) &= \mathcal{W}(\hat{\mathbf{x}}_k | \mathcal{Y}^{k-1}) \\ &= \int \mathcal{W}(\hat{\mathbf{x}}_k | \hat{\mathbf{x}}_{k-1}) \mathcal{W}(\hat{\mathbf{x}}_{k-1} | \mathcal{Y}^{k-1}) d\hat{\mathbf{x}}_{k-1}, \end{aligned} \tag{15}$$

which is the integral form of the FPE. In other words, a transient FPE solver is required to obtain the prior prediction $\mathcal{W}(\hat{\mathbf{x}}_k | \mathcal{Y}^{k-1})$ (i.e., the transient solution at the instance when the k^{th} measurement comes) given the posterior $\mathcal{W}(\hat{\mathbf{x}}_{k-1} | \mathcal{Y}^{k-1})$ as the initial PDF.

To facilitate the prediction step (especially for high-dimensional cases), the efficient transient FPE solver proposed in Refs.[19], [20] is adopted in current work. The transient solution $\mathcal{W}(t, \mathbf{x})$ of FPE in Eq.4 is sought in the following CANDECOMP/PARAFAC decomposition (CPD) form:

$$\mathcal{W}(t,\mathbf{x}) \approx \mathcal{U}(t,\mathbf{x}) = \sum_{l=1}^{R_U} \left[\left(\prod_{d=1}^{P} u_d^l(x_d) \right) T^l(t) \right], \quad (16)$$

where $u_d^l(x_d)$ and $T^l(t)$ for $l = 1, 2, \ldots, R_U$ are called spatial and temporal basis functions respectively. Note that each of these basis functions has only one variable and can be discretized by Chebyshev spectral method on the 1D domain. This results in vectors u_d^l and T^l with size $n_d \times 1$ and $n_t \times 1$ respectively. And for convenience, the discretized version of $\mathcal{U}$ is denoted by $\mathbb{U} = \sum_{l=1}^{R_U} \bigotimes_{d=1}^{P+1} f_d^l$, where f_d^l includes both u_d^l and T^l. Correspondingly, the Fokker Planck (FP) operator $\mathcal{L}_{\mathcal{FP}}$ can be written in the tensorized form as

$$\mathbb{A} = \sum_{i_A=1}^{R_A} \left[\left(\bigotimes_{d=1}^{P} A_d^{i_A} \right) \otimes I_t \right], \quad (17)$$

where $A_d^{i_A}$ are $n_d \times n_d$ matrices, and I_t is an $n_t \times n_t$ identity matrix. Interested readers can find more details in Ref.[19] on how this is achieved. Next the operator $\frac{\partial}{\partial t}$ in Eq.4 can be approximated as

$$\frac{\partial}{\partial t} \approx \left(\bigotimes_{d=1}^{P} I_d \right) \otimes D_t, \quad (18)$$

where I_d is an $n_d \times n_d$ identity matrix and D_t is the $n_t \times n_t$ first order differentiation matrix corresponding to the temporal domain. Now the FPE is reduced to

$$\mathbb{A}'\mathbb{U} = 0, \quad (19)$$

where $\mathbb{A}'$ is the sum of $\mathbb{A}$ and operator $\frac{\partial}{\partial t}$ in Eq.18, and is rewritten as $\mathbb{A}' = \sum_{i_A=1}^{R_{A'}} \bigotimes_{d=1}^{P+1} G_d^{i_A}$ for convenience.

Eq.19 can be solved by transforming into an optimization problem:

$$\min_{\{f_d^l\}} \parallel \mathbb{A}'\mathbb{U} \parallel_F^2, \quad (20)$$

and define $\mathcal{R} = \parallel \mathbb{A}'\mathbb{U} \parallel_F^2$. The necessary condition for minimization is

$$\frac{\partial \langle \mathbb{A}'\mathbb{U}, \mathbb{A}'\mathbb{U} \rangle}{\partial f_d^l} = 0, \quad (21)$$

for $d = 1, 2, \ldots, P+1$ and $l = 1, 2, \ldots, R_U$. Collecting terms in Eq.21 for all l's and fixed dimension d, we have

$$\underbrace{\begin{pmatrix} \mathbf{M}_{1,1} & \cdots & \mathbf{M}_{1,R_U} \\ \vdots & \ddots & \vdots \\ \mathbf{M}_{R_U,1} & \cdots & \mathbf{M}_{R_U,R_U} \end{pmatrix}}_{=\mathbf{M}} \begin{pmatrix} f_d^1 \\ \vdots \\ f_d^{R_U} \end{pmatrix} = 0, \quad (22)$$

where $\mathbf{M}_{i,j}$ is submatrix of the block matrix $\mathbf{M}$, given by

$$\mathbf{M}_{i,j} = \sum_{i_A=1}^{R_{A'}} \sum_{j_A=1}^{R_{A'}} (G_d^{j_A})^T G_d^{i_A} \prod_{k \neq d} \langle G_k^{i_A} f_k^j, G_k^{j_A} f_k^i \rangle \quad (23)$$

In the alternating least squares (ALS) framework [25], [19], Eq.22 is reduced to a linear system and solved sequentially for each dimension in an iterative manner. Therefore, the number of unknowns in a single iteration is independent of the dimensionality P once the user prescribes R_U. Moreover, a significant amount of computation is saved by noting that only a small portion of the terms involved need to be recalculated for different d and R_U. In implementation, we begin with $R_U = 1$ and random initial values for f_d^l, and then increase R_U gradually until stopping criteria are met. In order for the ALS scheme to return a nontrivial answer, we must incorporate the initial condition, boundary condition as described in Ref.[20].

For the update step, the Bayes' rule can be used as follows:

$$\begin{aligned} \mathcal{W}(\hat{\mathbf{x}}_k^+) = \mathcal{W}(\hat{\mathbf{x}}_k|\mathcal{Y}^k) &= \frac{\mathcal{W}(\hat{\mathbf{x}}_k, \mathbf{y}_k, \mathcal{Y}^{k-1})}{\mathcal{W}(\mathbf{y}_k, \mathcal{Y}^{k-1})} \\ &= \frac{\mathcal{W}(\hat{\mathbf{x}}_k, \mathbf{y}_k|\mathcal{Y}^{k-1})\mathcal{W}(\mathcal{Y}^{k-1})}{\mathcal{W}(\mathbf{y}_k|\mathcal{Y}^{k-1})\mathcal{W}(\mathcal{Y}^{k-1})} \\ &= \frac{\mathcal{W}(\hat{\mathbf{x}}_k, \mathbf{y}_k|\mathcal{Y}^{k-1})}{\mathcal{W}(\mathbf{y}_k|\mathcal{Y}^{k-1})} \\ &= \frac{\mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k, \mathcal{Y}^{k-1})\mathcal{W}(\hat{\mathbf{x}}_k|\mathcal{Y}^{k-1})}{\int \mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k, \mathcal{Y}^{k-1})\mathcal{W}(\hat{\mathbf{x}}_k|\mathcal{Y}^{k-1}) d\hat{\mathbf{x}}_k} \\ &= \frac{\mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k)\mathcal{W}(\hat{\mathbf{x}}_k|\mathcal{Y}^{k-1})}{\int \mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k)\mathcal{W}(\hat{\mathbf{x}}_k|\mathcal{Y}^{k-1}) d\hat{\mathbf{x}}_k} \\ &= \frac{\mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k)\mathcal{W}(\hat{\mathbf{x}}_k^-)}{\int \mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k)\mathcal{W}(\hat{\mathbf{x}}_k^-) d\hat{\mathbf{x}}_k} \end{aligned} \quad (24)$$

where the last but one step uses the fact that $\mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k, \mathcal{Y}^{k-1}) = \mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k)$ since $\hat{\mathbf{x}}_k$ has already incorporated the information of $\mathcal{Y}^{k-1}$.

In Eq.24, the likelihood PDF $\mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k)$ is chosen as [9]:

$$\mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k) = \frac{\exp\left(-\frac{1}{2}[\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k)]^T \mathbf{R}^{-1} [\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k)]\right)}{\sqrt{(2\pi)^Q \det(\mathbf{R})}}, \quad (25)$$

and the denominator of Eq.24 is essentially a normalization constant to ensure that $\mathcal{W}(\hat{\mathbf{x}}_k^+)$ is a valid PDF. Since the prior $\mathcal{W}(\hat{\mathbf{x}}_k^-)$ is given by the tensor solver, it is in the CPD form, i.e., all the independent variables are decoupled. Thus if the likelihood $\mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k)$ is also in the CPD form, the muti-dimensional integration in Eq.24 will be trivial. When this is not the case, $\mathcal{W}(\mathbf{y}_k|\hat{\mathbf{x}}_k)$ can be approximated in the CPD form by using the ALS algorithm described for example in Ref.[19].

V. Numerical Examples

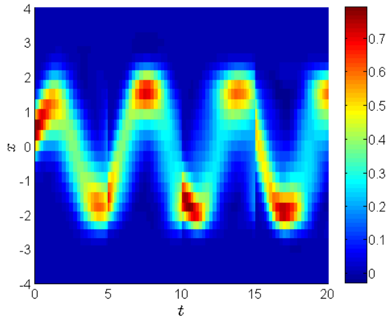
In this section, 2 examples of $2D$ state space are examined, where the results by the FPE based Bayesian filter are compared with those of the extended Kalman filer.

A. Results for $2D$ System 1

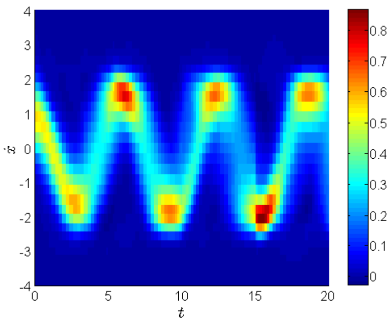
Consider the following nonlinear oscillator with $2D$ state space:

$$\ddot{x} + b\dot{x} + x + a(x^2 + \dot{x}^2)\dot{x} = g\xi(t), \quad (26)$$

where $\xi(t)$ is Gaussian white noise with intensity $\mathbf{Q}$. For simulation of the true model, let the parameters have the following values: $a = 0.125, b = -0.5, g = 1, \mathbf{Q} = 0.4$ and the true initial condition is $[0, 0]^T$. While in the filters,



(a) For state x



(b) For state $\dot{x}$

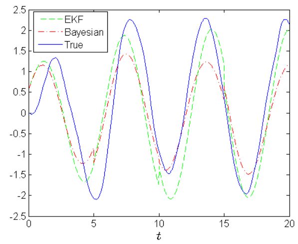
Fig. 1. $1D$ marginal state PDF by FPE based Bayesian filter for system 1

we assume the model parameters are given by $a = 0.1, b = -0.4, g = 1, \mathbf{Q} = 0.2$ and use $[1, 1]^T, \mathrm{diag}([0.5, 0.5])$ as the mean and covariance of the initial condition respectively. Suppose the measurement model is given by

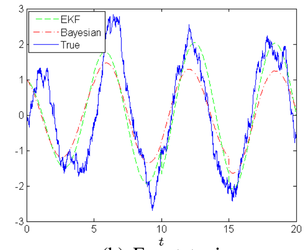
$$\mathbf{y}_k = x_k + \mathbf{v}_k, \tag{27}$$

i.e., only position x is observed. The measurement arrives after $5s$ intervals with measurement noise covariance $\mathbf{R} = 0.5$. The Milstein's stochastic integration scheme [26] is used to propagate the true model and the model in EKF. For the Bayesian filter, 20 enrichment steps are performed by the tensor solver for each propagation stage.

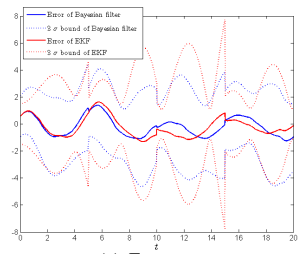
The evolution of the prior and posterior conditional $2D$ state PDFs is given in Figure 3 and the $1D$ marginal state PDFs are shown in Figure 1. As can be seen, the state PDF is clearly non-Gaussian following the propagation stage. When a measurements arrives, in the update step the PDF "shrinks" due to gain of information but still remains non-Gaussian. These conditional state PDFs contain the full probabilistic information of the system and can be used for calculating desired expectations of the state and other quantities. For state estimation, the mean and variance of the $1D$ marginal state PDFs can be obtained as shown in Figure 1, and based on which the state estimate is compared with that of EKF as illustrated in Figures 2(a) and 2(b). The corresponding error and 3σ error bounds comparison are given in Figures 2(c) and 2(d), which show that the state estimate errors are comparable, but the nonlinear Bayesian filter has tighter error bounds, at the cost of more computational effort.



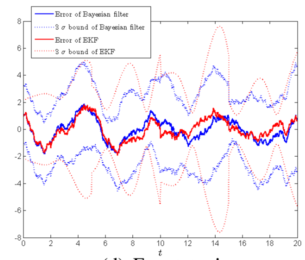
(a) For state x



(b) For state $\dot{x}$



(c) For state x



(d) For state $\dot{x}$

Fig. 2. Results comparison for FPE based Bayesian filter and EKF for system 1

B. *Results for $2D$ System 2*

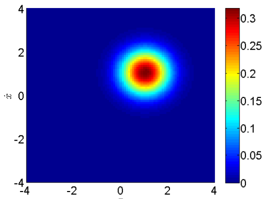
Next consider the Van der Pol oscillator given by

$$\ddot{x} + v_1 x + v_2 \dot{x}(1 - x^2) = g\xi, \tag{28}$$

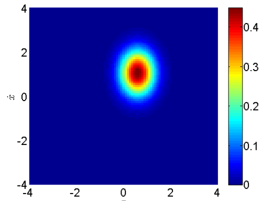
where $\xi(t)$ is Gaussian white noise with intensity $\mathbf{Q}$. This nonlinear system is characterized by its bimodal limit cycle in the $2D$ state PDF. Let the true model have the following parameters: $v_1 = 1, v_2 = -1, g = 1, \mathbf{Q} = 1$ and the true initial condition is $[0, 0]^T$. We assume the model parameters used in EKF and the Bayesian filter are given by $v_1 = 2, v_2 = -1, g = 1, \mathbf{Q} = 4$ and use $[1, 1]^T, \mathrm{diag}([1, 1])$ as the mean and covariance of the initial condition respectively. Suppose the same measurement model in Eq.27 are chosen with $\mathbf{R} = 0.5$ and measurement is available at $4s$ intervals.

The evolution of the prior and posterior conditional $2D$ state PDFs as well as the $1D$ marginal state PDFs are given in Figure 6 and Figure 4 respectively. The limit cycle with its double peaks is clearly visible at the end of each propagation stage since the $4s$ interval is long enough for the state PDF to reach the stationary state. When the measurement becomes available, the posterior PDF "sheds" one of the two peaks of the prior PDF. However, with the poor knowledge of system parameters and sparse measurement data, none of these two filters perform well enough in terms of state estimation as shown in Figure 5. As can be seen in Figures 5(c) and 5(d), the EKF state estimate has larger error (with MSE $= 6.34$ for x and 10.27 for $\dot{x}$) than the Bayesian filter (with MSE $= 1.32$ for x and 2.16 for $\dot{x}$) and the error bounds of EKF become inconsistent.

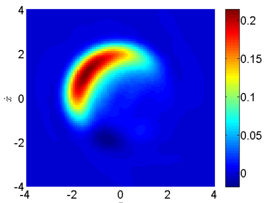
This example illustrates that using mean of the conditional



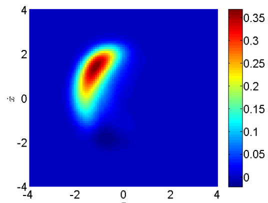
(a) Prior PDF at $t = 0s$



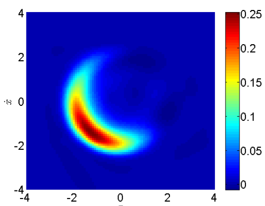
(b) Posterior PDF at $t = 0s$



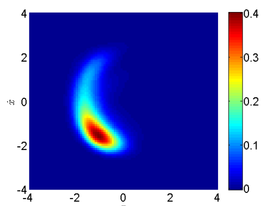
(c) Prior PDF at $t = 5s$



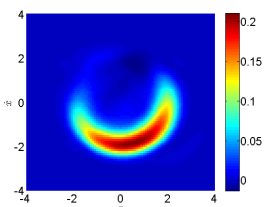
(d) Posterior PDF at $t = 5s$



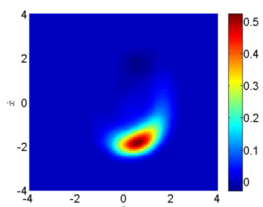
(e) Prior PDF at $t = 10s$



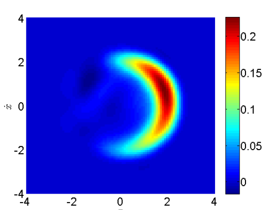
(f) Posterior PDF at $t = 10s$



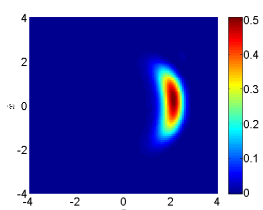
(g) Prior PDF at $t = 15s$



(h) Posterior PDF at $t = 15s$



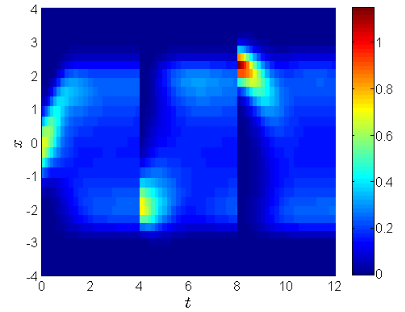
(i) Prior PDF at $t = 20s$



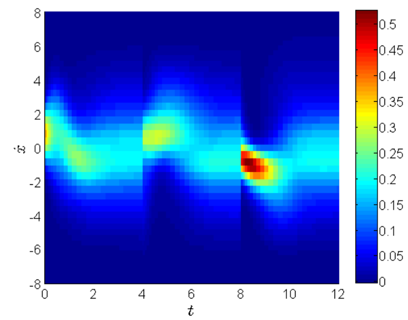
(j) Posterior PDF at $t = 20s$

Fig. 3. $2D$ state PDF by FPE based Bayesian filter for system 1

state PDF as state estimate can be highly inadequate. In this example, the symmetric bimodal nature of the state PDF (see Figure 6 and 4) causes the mean to be zero and therefore is clearly a poor state estimate. On the other hand, the nonlinear Bayesian filter provides the full PDF, i.e., a better estimate of the state in terms of its probability distribution over the entire state space. Moreover, it should be aware that only one simulation of the true model is performed, and based on which the error of the two filters are computed. If more simulations are performed, by definition the average state trajectories will undoubtedly become closer to the FPE based nonlinear Bayesian filter results.

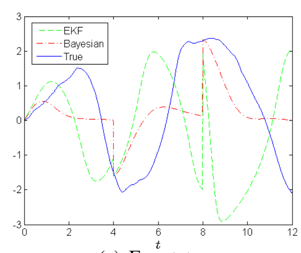


(a) For state x

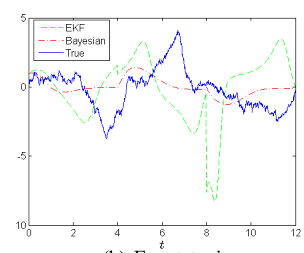


(b) For state $\dot{x}$

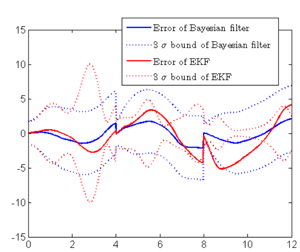
Fig. 4. $1D$ marginal state PDF by FPE based Bayesian filter for system 2



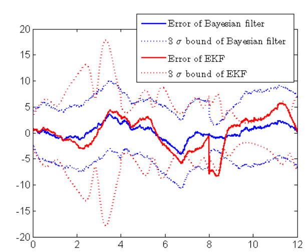
(a) For state x



(b) For state $\dot{x}$



(c) For state x

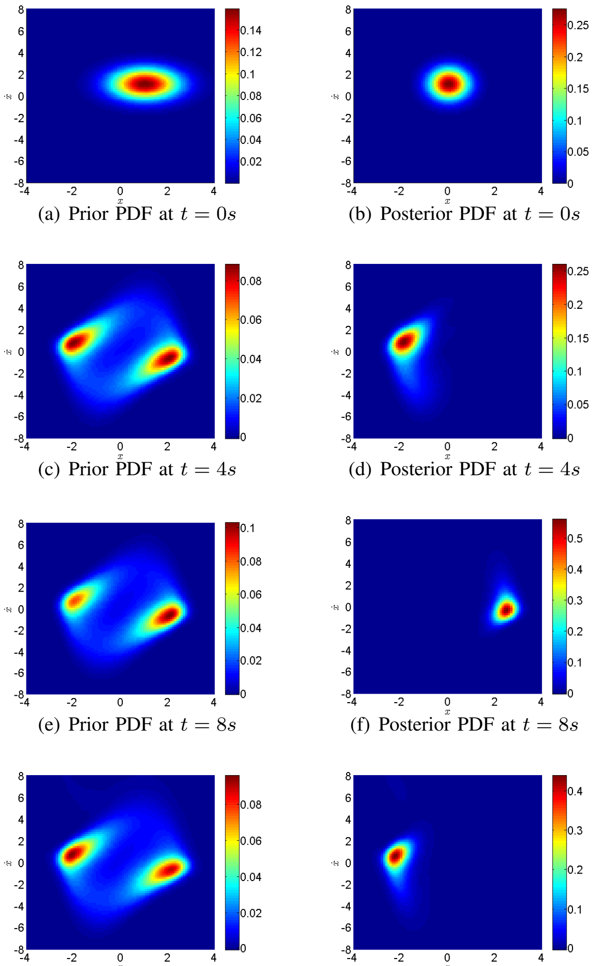


(d) For state $\dot{x}$

Fig. 5. Results comparison for FPE based Bayesian filter and EKF for system 2

VI. Conclusions

In this paper, the a nonlinear Bayesian filter is developed by using the transient tensor solver proposed in Refs.[19],



(a) Prior PDF at $t = 0s$ (b) Posterior PDF at $t = 0s$

(c) Prior PDF at $t = 4s$ (d) Posterior PDF at $t = 4s$

(e) Prior PDF at $t = 8s$ (f) Posterior PDF at $t = 8s$

(g) Prior PDF at $t = 12s$ (h) Posterior PDF at $t = 12s$

Fig. 6. $2D$ state PDF by FPE based Bayesian filter for system 2

[20]. The nonlinear, PDF based approach illustrates the pitfalls of using the "mean" as state estimate (especially in multi-modal cases) and provides a wholistic view of the system's whereabouts. The state estimate results are compared with those of the EKF showing that the Bayesian filter has better error bounds for system 1 of section V-A and more accurate state estimate in system 2 of section V-B at the cost of more computational effort.

REFERENCES

[1] P. S. Maybeck, Stochastic models, estimation, and control. Academic press, 1982, vol. 3.

[2] N. J. Gordon, D. J. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in IEE Proceedings F (Radar and Signal Processing), vol. 140, no. 2. IET, 1993, pp. 107–113.

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

[4] I. B. Rhodes, "A tutorial introduction to estimation and filtering," Automatic Control, IEEE Transactions on, vol. 16, no. 6, pp. 688–706, 1971.

[5] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," Automatic Control, IEEE Transactions on, vol. 45, no. 5, pp. 910–927, 2000.

[6] J. L. Crassidis and J. L. Junkins, Optimal estimation of dynamic systems. CRC press, 2011.

[7] S. Challa and Y. Bar-Shalom, "Nonlinear filter design using Fokker-Planck-Kolmogorov probability density evolutions," Aerospace and Electronic Systems, IEEE Transactions on, vol. 36, no. 1, pp. 309–315, 2000.

[8] J. Yoon, Nonlinear Bayesian estimation via solution of the Fokker-Planck equation, 2009.

[9] M. Kumar and S. Chakravorty, "Nonlinear filter based on the Fokker-Planck equation," Journal of Guidance, Control, and Dynamics, vol. 35, no. 1, pp. 68–79, 2012.

[10] L. Li, R. Mei, and J. F. Klausner, "Boundary conditions for thermal lattice boltzmann equation method," Journal of Computational Physics, vol. 237, pp. 366–395, 2013.

[11] ——, "Heat transfer evaluation on curved boundaries in thermal lattice boltzmann equation method," Journal of Heat Transfer, vol. 136, no. 1, p. 012403, 2014.

[12] R. E. Bellman, Dynamic programming. Princeton University Press, Princeton, New Jersey, 1957.

[13] Y. Sun and M. Kumar, "A Markov chain Monte Carlo particle solution of the initial uncertainty propagation problem," in AIAA Guidance, Navigation, and Control Conference, 2012.

[14] M. Kumar, Y. Sun, and S. Thakur, "On the stochastic modeling of subcritical primary atomization in liquid rocket engines," in Science and Technology Forum and Exposition. National Harbor, MD: AIAA, Jan. 13-17, 2014.

[15] C. Yang and M. Kumar, "An evaluation of Monte Carlo for nonlinear initial uncertainty propagation in Keplerian mechanics," in International Conference on Information Fusion. Washington, D.C.: IEEE, 2015.

[16] M. Kumar, S. Chakravorty, P. Singla, and J. L. Junkins, "The partition of unity finite element approach with hp-refinement for the stationary Fokker-Planck equation," Journal of Sound and Vibration, vol. 327, no. 1, pp. 144–162, 2009.

[17] M. Kumar, S. Chakravorty, and J. L. Junkins, "A semianalytic meshless approach to the transient Fokker-Planck equation," Probabilistic Engineering Mechanics, vol. 25, no. 3, pp. 323–331, 2010.

[18] Y. Sun and M. Kumar, "A meshless p-PUFEM Fokker-Planck equation solver with automatic boundary condition enforcement," in American Control Conference (ACC), 2012. IEEE, 2012, pp. 74–79.

[19] ——, "Numerical solution of high dimensional stationary Fokker-Planck equations via tensor decomposition and Chebyshev spectral differentiation," Computers & Mathematics with Applications, vol. 67, no. 10, pp. 1960–1977, 2014.

[20] ——, "A numerical solver for high dimensional transient Fokker-Planck equation in modeling polymeric fluids," Journal of Computational Physics, vol. 289, pp. 149–168, 2015.

[21] ——, "A tensor decomposition approach to high dimensional stationary Fokker-Planck equations," in American Control Conference (ACC), 2014. Portland, OR: IEEE, Jun. 4-6, 2014.

[22] ——, "Solution of high dimensional transient Fokker-Planck equations by tensor decomposition," in American Control Conference (ACC), 2015. Chicago, IL: IEEE, July 1-3, 2015.

[23] ——, "A tensor decomposition method for high dimensional Fokker-Planck equation modeling polymeric liquids," in Science and Technology Forum and Exposition. Kissimmee, FL: AIAA, Jan. 5-9, 2015.

[24] ——, "Uncertainty propagation in orbital mechanics via tensor decomposition," Celestial Mechanics and Dynamical Astronomy, under review.

[25] G. Beylkin and M. Mohlenkamp, "Algorithms for numerical analysis in high dimensions," SIAM Journal on Scientific Computing, vol. 26, no. 6, pp. 2133–2159, 2005.

[26] G. N. Milstein, Numerical integration of stochastic differential equations. Springer, 1995.