Solving Nonlinear Filtering Problems in Real Time by Legendre Galerkin Spectral Method

Wenhui Dong, Xue Luo, *Senior Member, IEEE*, and Stephen S.-T. Yau, *Fellow, IEEE*

Dedicated to Prof. R. W. Brockett on the occasion of his 82nd birthday

***Abstract*—It is well known that the nonlinear filtering (NLF) problem has important applications in both military and civil industries. The central question is to solve the posterior conditional density function of the states, which satisfies the Kushner or the Duncan–Mortensen–Zakai (DMZ) equation after suitable change of probability measure. In this article, we shall follow the so-called Yau–Yau's algorithm to split the solution of the DMZ equation into on- and off-line part, where the off-line part is to solve the forward Kolmogorov equation (FKE) with the initial conditions to be the orthonormal bases in some suitable function space. Instead of the generalized Hermite function investigated by the second and the third author of this article, we shall explore the generalized Legendre polynomials. The Legendre spectral method (LSM) is used to numerically solve the FKE. Under certain conditions, the convergence rate of LSM is twice faster than that of the Hermite spectral method. Two two-dimensional numerical experiments of NLF problems (time-invariant and time-varying cases) have been numerically solved to illustrate the feasibility of our algorithm. Our algorithm outperforms the extended Kalman Filter and particle filter in both real-time manner and accuracy.**

***Index Terms*—Convergence analysis, forward Kolmogorov equation (FKE), Legendre spectral method (LSM), nonlinear filtering (NLF).**

Manuscript received November 1, 2019; accepted May 15, 2020. Date of publication June 17, 2020; date of current version March 29, 2021. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 11961141005, in part by the Tsinghua University start-up fund, and in part by the Tsinghua University Education Foundation fund under Grant 042202008. The work of Xue Luo was supported in part by the NSFC under Grant 11871003 and in part by the Fundamental Research Funds for the Central Universities under Grant YWF-20-BJ-J-644. Recommended by Associate Editor G. Gu. *(Corresponding author: Stephen S.-T. Yau.)*

Wenhui Dong was with the Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China. She is now with the School of Mathematics and Information Science, Shaanxi Normal University, Xian 710062, China (e-mail: dwh15@mails.tsinghua.edu.cn).

Xue Luo is with the School of Mathematical Sciences, Beihang University, Beijing 102206, China (e-mail: xluo@buaa.edu.cn).

Stephen S.-T. Yau is with the Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China (e-mail: yau@uic.edu).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2020.3002979

I. Introduction

THE field of nonlinear filtering (NLF) has its origin from tracking and signal processing problems. The states in the stochastic dynamical systems are also called signals. They represent all kinds of quantities in various applications. It is extensively noticed that the goal of NLF is to obtain best estimate of the states, recursively in time, based on the noisy observations of the states. In many real applications, say the tracking problems, the real-time performance is highly appreciated. Research in filtering problem can be dated back to almost two centuries to the work of Gauss. Later, the famous control scientist Wiener advanced the filtering theory. In the early 1960s, two most influential papers [20], [21] were published in the ASME Journal of Basic Engineering, in which the so-called Kalman filter (KF) and Kalman–Bucy filter were derived for the first time. Despite its success in many practical applications, the limitations on the nonlinearity and non-Gaussian assumption of the initial probability density of the KF spurred and pushed the mathematicians and scientists to seek for the optimal NLF. Meanwhile, since most stochastic dynamic systems considered in various practical applications are nonlinear, one direction is to modify KF to adapt the nonlinearities. The researchers developed extended KF (EKF), unscented KF (UKF)[19], ensemble KF (EnKF) [13], etc., which can tackle with weak nonlinearities (that is almost linear). Essentially, EKF, which is the simplest filter for NLF systems, performs poorly when the dynamic system is significantly nonlinear and is very sensitive to initial value due to Taylor approximation, and even may completely fail (see Figs. 1 and 2). UKF assumes the posterior distribution of the state is Gaussian which also restricts its applications. EnKF is the NLF algorithm integrating the data assimilation into ensemble generation problem, and has been the key ingredient of prediction and predictability research for weather and oceanic prediction applications [3], [4], [26].

The other direction of designing optimal NLF is the particle filter (PF), referring to such as [5] and [8], which is developed from sequential Monte Carlo method and becomes the most popular method nowadays. The PF is applicable to nonlinear, non-Gaussian state update and observation equations, and can become asymptotically optimal as the number of particles goes to infinity. However, the main drawback of this method is that it is hard to be implemented as a real-time application, due to its essence of Monto Carlo simulation. "Real-time" means that the estimation of the states is made on the spot instantaneously,

while the observation data keep coming in. Hence, it is necessary to develop a real-time solver to the NLF problems.

In the 1960s, Duncan[10], Mortensen [31], and Zakai [49] independently derived the so-called Duncan–Mortensen–Zakai's (DMZ) equation, a stochastic partial differential equation (SPDE), which is satisfied by the unnormalized conditional density function of the states. It is well known that the exact solution to the DMZ equation, generally speaking, can not be written in a closed form. With the promotion of the computational power, many mathematicians make effort to seek an efficient algorithm to construct a "good" approximate solution to the DMZ equation. Although the numerical methods to SPDE can yield the approximate solution, it is inevitable to burden the heavy and intensive computations. One of the methods to relieve the on-line computational load is the splitting-up method originated from the Trotter product formula [6], [7], [27], and [48]. In [28], the second and the third author of this article generalized Yau–Yau's algorithm [48] to the time-varying settings of the NLF problems, where the drift term, diffusion term, and observation term in (1) could be explicitly time-dependent.

It has been a long history of using the spectral method with orthogonal polynomials to solve various problems, such as in the field of computational fluid dynamics, which can be dated back to 1970s [15]. It provides very accurate approximation with a relatively small number of modes if the solution is smooth, (see [9]). Actually, the second and third author of this article made on- and off-line algorithm feasible by using Hermite spectral method (HSM) to solve the FKE [29]. The selection of Hermite functions is that they are defined on the unbounded domain decaying exponentially at infinity without worrying about the boundary conditions. However, they found that the NLF problems can be solved accurately by on- and off-line algorithm using HSM, only when some key parameters are properly tuned. Although the generalized Jacobi polynomials are also introduced and investigated to numerically solve the FKE arising from NLF problems in [30], they are only used in the scalar NLF problem. As it is known that the solution of the FKE can actually be considered as the probability density function (PDF), which is usually assumed to vanish at infinity and satisfy the normality constraint and positivity constraint, thus it leads us to find the solution of the FKE in the bounded domain with zero-boundary. In this article, we propose to solve the FKE of general dimensions in a bounded domain with Dirichlet boundary condition by using Legendre polynomials, see the numerical Section IV, as similar as the off-line ingredient of the algorithm in [29].

In the literature, a large number of spectral methods are discussed with a wide range of variations where the boundary conditions are enforced. In [35], Chebyschev approximations are employed to solve the 1-D FKE in the presence of two barriers a finite distance apart, and further the solutions are presented for the fundamental intervals $(-1, +1)$ and $(0, +1)$. The eigenfunction expansions of FKE for the first and second order nonlinear systems are discussed in [2], [17], [18], and [36]. In [25], by taking Fourier transformation in FKE, the analytic solutions of FKE with special initial conditions were obtained in the linear filtering system. In [45], with the same transformation in FKE, then it becomes an initial-boundary value problem about the characteristic function of the states, which is solved by using the finite difference (FD) technique. Later on, with the same transformation onto the FKE first, high-order FD schemes to solve the FKE have been developed [44].

Back to the real applications and computer simulations, then intuitively, the states to be estimated in the stochastic dynamical systems are normally in a large enough but bounded domain, thus the solution of the FKE is posed in a cylinder $[-a, a]^d \times [0, \infty]$ likewise, where d is the dimension of the states. It is then a question how to include the zero-boundary condition naturally in the orthogonal polynomials on a bounded domain. It is motivated by Shen [38] that the zero-boundary Legendre polynomials can be constructed from the classical ones, where the associated spectral method in solving second- and forth-order elliptic PDE have also been investigated. In this article, we shall study the zero-boundary Legendre spectral method (LSM) to solve the FKE served as the off-line part of Yau–Yau's algorithm, which is closely related to the implementation of the algorithm developed in [28], or (see Appendix A, [29]). Under certain conditions, the convergence rate of LSM is twice faster than the HSM. Two 2-D numerical experiments of NLF problems, including time-invariant and time-varying cases, have been numerically solved to illustrate the feasibility of our algorithm. Our algorithm outperforms the EKF and PF in both real-time manner and accuracy.

In order to maintain accuracy in traditional discretization-based numerical methods such as the finite element (FE) and FD method, it has been widely observed that the degrees of freedom of the approximation, i.e., the number of unknowns grow exponentially as the dimensionality of the underlying state space increases. This well-known curse of dimensionality fundamentally limits the use of FKE for NLF in high dimensional systems. Extensive research works have been devoted to developing efficient numerical solvers for high dimensional FKE [23], [24], [39]. In [46] and [47], the authors have done the numerical integration of high dimensional FKE by using the generalized Laguerre polynomials. They have previously applied standard Bubnov–Galerkin FE techniques to solve the FKE in two and three dimensions [37], [43]. Dolgov *et al.* [12] used tensor train format [32], [33] for parabolic PDEs in modeling polymeric liquids where the FKE of the states up to 12 dimensions was successfully solved. Sun and Kumar [40], [41] utilized tensor decomposition approach combined with Chebyshev spectral differentiation to tackle the FKE. The generalized LSM inevitably encounters the curse of dimensionality when the state dimension d grows, and the medium or low dimensional NLF problems are successfully solved in this article. Numerically solving high dimensional NLF problems by using generalized LSM is one of our future research topics.

There are many research works about using Galerkin spectral methods to solve the unconstrained/constrained optimal control problems, such as [14] and [16]. Recently in [42], the adaptive fuzzy control problem has been investigated for a class of nontriangular structural stochastic switched nonlinear systems with full state constraints. In [34], a fuzzy adaptive event-triggered control strategy was designed for pure-feedback

nonlinear systems as a more general class of triangular structural systems. One of our future research focuses is how to combine the Galerkin spectral method with orthogonal polynomials to solve this kind of fuzzy control problem, even extending it to the complex nonlinear systems with constraint states or prescribed performance.

This article is organized as follows. Section II gives some preliminaries, such as the general filtering problems, the on- and off-line algorithm in [28], [29], and the classical Legendre polynomials. The LSM is introduced in Section II-B. In Section III, we focuses on the analysis of the convergence rate of LSM to FKE. In Section IV, the detailed formulation of our algorithm is presented, as well as the numerical simulations of two 2-D NLF problems, including both time-invariant and time-varying examples. We arrive at our conclusions in Section V.

II. Preliminaries

In this section, we shall recall the general NLF problems and the on- and off-line algorithm introduced in [28] and [29]. The classical Legendre polynomials and its properties are also recalled.

A. On- and Off-Line Algorithm for the NLF Problems

In this article, we consider the following signal observation model:

$$\begin{cases} dx_t = f(x_t,t)dt + G(x_t,t)dv_t \\ dy_t = h(x_t,t)dt + dw_t \end{cases} \tag{1}$$

where x_t and f are d-vectors, G is a $d \times r$ matrix, and v_t is a r-vector Brownian motion process with $E[dv_t dv_t^\top] = Q(t)dt$, with $Q(t) \in \mathbb{R}^{r\times r}$. Besides, y_t and h are m-vectors and w_t is a Brownian motion process with $E[dw_t dw_t^\top] = S(t)dt$, with $S \in \mathbb{R}^{m\times m}$. x_t is the state of the system at time t with x_0 satisfying some distribution, and y_t is the observation at time t with $y_0 = 0$. We assume that v_t, w_t, and x_0 are mutually independent.

Let $p(x,t)$ denote the conditional probability density of the state x_t given the observation history $\{y_s : 0 \le s \le t\}$, which evolves according to the Kushner's equation [22]. Under suitable change of probability measure, $p(x,t)$ can be transformed to the unnormalized conditional density of the state x_t conditioned on the observation history $Y_t = \{y_s : 0 \le s \le t\}$, denoted as $\sigma(x,t)$, which satisfies the DMZ equation [10], [31], [49] as follows:

$$\begin{cases} d\sigma(x,t) = L\sigma(x,t)dt + \sigma(x,t)h^\top(x,t)S^{-1}(t)dy_t \\ \sigma(x,0) = \sigma_0(x) \end{cases} \tag{2}$$

where $\sigma_0(x)$ is the probability density of the initial state x_0, and

$$L(*) = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}\left[(GQG^\top)_{ij}*\right] - \sum_{i=1}^{d}\frac{\partial (f_i *)}{\partial x_i}. \tag{3}$$

The normalized conditional density $p(x,t)$ is then given by $p(x,t) = \frac{\sigma(x,t)}{\int \sigma(x,t)dx}$. Compared with the Kushner's equation, the DMZ equation (2) is much easier and cheaper to solve. In [28], the second and the third author of this article designed an on- and off-line algorithm for DMZ equation, which yields robust state estimation from a given sample path. Given an observation path y_t, through an invertible exponential transformation [11], one obtains that

$$\rho(x,t) = \exp\left[-h^\top(x,t)S^{-1}(t)y_t\right]\sigma(x,t) \tag{4}$$

where $\rho(x,t)$ satisfies the "pathwise-robust" DMZ equation which involves y_t only in the coefficients of the PDE

$$\begin{cases} \dfrac{\partial \rho}{\partial t}(x,t) + \dfrac{\partial}{\partial_t}\left(h^\top S^{-1}\right)^\top y_t \rho(x,t) \\ \qquad = \exp(-h^\top S^{-1} y_t)\left[L - \dfrac{1}{2}h^\top S^{-1}h\right] \\ \qquad \cdot \left[\exp(h^\top S^{-1}y_t)\rho(x,t)\right] \\ \rho(x,0) = \sigma_0(x). \end{cases} \tag{5}$$

We set up the observation time sequence as $\mathcal{P}_k = \{0 = \tau_0 \le \tau_1 \le \cdots \le \tau_k = T\}$. Let $\rho_i, i = 1,\ldots,k$ be the solution of the "pathwise-robust" DMZ equation (5) on the time interval $[\tau_{i-1},\tau_i]$, with y_t in (5) replaced by $y_{\tau_{i-1}}$

$$\begin{cases} \dfrac{\partial \rho_i}{\partial t}(x,t) + \dfrac{\partial}{\partial t}(h^\top S^{-1})^\top y_{\tau_{i-1}}\rho_i(x,t) \\ \qquad = \exp(-h^\top S^{-1}y_{\tau_{i-1}})\left[L - \dfrac{1}{2}h^\top S^{-1}h\right] \\ \qquad \cdot \left[\exp(h^\top S^{-1}y_{\tau_{i-1}})\rho_i(x,t)\right] \\ \rho_1(x,0) = \sigma_0(x) \\ \text{or} \\ \rho_i(x,\tau_{i-1}) = \rho_{i-1}(x,\tau_{i-1}), i = 2,\ldots,k-1. \end{cases} \tag{6}$$

Intuitively, we have ρ_i glued in time to yield an approximation to ρ in (5), i.e.,

$$\rho(x,t) \approx \sum_{i=1}^{k} \chi_{[\tau_{i-1},\tau_i]}(t)\rho_i(x,t) \tag{7}$$

where $\chi_{[a,b]}(t)$ is the indicator function on the interval $[a,b]$, i.e., $\chi_{[a,b]}(t) = 1$, if $t \in [a,b]$ or 0, otherwise. After a similar exponential transformation (4) back, the observation disappears in the evolution equation, but just in the initial update. This is the key ingredient of the algorithm in [48], so is in ours.

Proposition 1: (Prop. 2.1, [28]) For each $\tau_{i-1} \le t < \tau_i, i = 1,2,\ldots,k$, $\rho_i(x,t)$ satisfies (5) with $y_t = y_{\tau_{i-1}}$ if and only if

$$u_i(x,t) = \exp\left[h^\top(x,t)S^{-1}(t)y_{\tau_{i-1}}\right]\rho_i(x,t) \tag{8}$$

satisfies the FKE

$$\frac{\partial u_i}{\partial t}(x,t) = \left(L - \frac{1}{2}h^\top S^{-1}h\right)u_i(x,t) \tag{9}$$

Algorithm 1: On-Line Algorithm.

1: **Initialization:** given $\Omega, T, \triangle t, y_0 = 0, \sigma_0(x)$,
Let $k = \frac{T}{\triangle t}$, and $\{0 = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_k = T\}$.
Let $u_1(x,0) = \sigma_0(x)$, and normalize it as
$u_1(x,0) = \frac{u_1(x,0)}{\int_\Omega u_1(x,0)dx}$.
2: Given $u_1(x,0)$, we obtain $u_1(x,\tau_1)$ by Algorithm 2.
3: At time instant $t = \tau_1$, when the new observation y_{τ_1} is available, we obtain
$u_2(x,\tau_1) = exp[h^T(x,\tau_1)S^{-1}(\tau_1)y_{\tau_1}]u_1(x,\tau_1)$,
then normalize it as $u_2(x,\tau_1) = \frac{u_2(x,\tau_1)}{\int_\Omega u_2(x,\tau_1)dx}$.
4: **for** $i = 2$ to k **do**
5: $u_i(x,\tau_{i-1}) =$
$\exp[h^T(x,\tau_{i-1})S^{-1}(\tau_{i-1})(y_{\tau_{i-1}} - y_{\tau_{i-2}})]u_{i-1}(x,\tau_{i-1})$,
where $u_{i-1}(x,\tau_{i-1})$ is from Algorithm 2.
6: **end for**

Algorithm 2: Off-Line Algorithm.

1: **Initialization:** given $u_1(x,\tau_0)$ in Algorithm 1.
2: **for** $i = 1$ to k **do**
3: Solve the FKE (9) to get $u_i(x,t)$.
4: Let $t = \tau_i$ in $u_i(x,t)$, then get $u_i(x,\tau_i)$.
Normalize $u_i(x,\tau_i) = \frac{u_i(x,\tau_i)}{\int_\Omega u_i(x,\tau_i)dx}$.
Obtain $\rho_i(x,\tau_i)$ by (8) and the normalized $u_i(x,\tau_i)$.
Normalize $\rho_i(x,\tau_i) = \frac{\rho_i(x,\tau_i)}{\int_\Omega \rho_i(x,\tau_i)dx}$.
5: **end for**

where L is defined in (3), while the initial data is updated as follows:

$$\begin{cases} u_1(x,0) = \sigma_0(x) \\ \qquad \text{or} \\ u_i(x,\tau_{i-1}) = \exp\left[h^\top(x,\tau_{i-1})S^{-1}(\tau_{i-1})(y_{\tau_{i-1}} - y_{\tau_{i-2}})\right] \\ \qquad\qquad \cdot u_{i-1}(x,\tau_{i-1}), \quad i \geq 2. \end{cases} \tag{10}$$

The on- and off-line algorithm for NLF problems are described in Algorithms 1 and 2, respectively.

B. Legendre Galerkin Spectral Approximation

Let us first introduce some basic notations and properties of classical univariate Legendre polynomials. Define $L^2([-1,1])$ be the Lebesgue space, equipped with the inner product $\langle f,g\rangle = \int_{-1}^{1} f(x)g(x)dx$, for any $f,g \in L^2([-1,1])$, and the induced norm is $||f||^2 = \langle f,f\rangle$. The well-known Legendre polynomial of degree n denoted as $L_n(x)$, is given by the recurrence formula as follow:

$$\begin{aligned} L_0(x) &= 1 \\ L_1(x) &= x \\ L_{n+1}(x) &= \frac{2n+1}{n+1}xL_n(x) - \frac{n}{n+1}L_{n-1}(x) \end{aligned} \tag{11}$$

for $n = 1,2,\ldots, x \in [-1,1]$. Let us denote

$$\begin{aligned} S_N &:= \text{span}\{L_0(x), L_1(x), \ldots, L_N(x)\} \\ V_N &:= \{v \in S_N : v(\pm 1) = 0\}. \end{aligned}$$

It is clear to see that $V_N \subset S_N$.

One of the useful facts of Legendre polynomials is that they are mutually orthogonal with respect to the weight function $w(x) = 1$ and $x \in [-1,1]$, i.e.,

$$\langle L_k(x), L_j(x)\rangle = \begin{cases} \frac{2}{2k+1}, & \text{if } j = k \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

For the multivariate case, i.e., the domain $I^d = [-1,1]^d$, $d \geq 2$. Similarly, $L^2(I^d)$ denotes the square integrable functions on I^d, equipped with the inner product $\langle u,v\rangle = \int_{I^d} uvd\vec{\mathbf{x}}$, with $\vec{\mathbf{x}} = (x_1,\ldots,x_d) \in I^d \subset \mathbb{R}^d$ and the induced norm $||u||^2 = \langle u,u\rangle$, for any $u,v \in L^2(I^d)$. Let us denote S_N^d the space of multivariate Legendre polynomials of degree up to N. The tensor product of the Legendre polynomials

$$\mathbf{L}_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) := L_{k_1}(x_1)\cdots L_{k_d}(x_d) \tag{13}$$

for $\vec{\mathbf{k}} = (k_1,\ldots,k_d)$, $k_i \in \mathbb{Z}_+$, $1 \leq i \leq d$ are considered as the orthogonal base in $L^2(I^d)$. Let $\mathbb{P}_N$ be the orthogonal projection operator from $L^2(I^d)$ upon S_N^d, i.e., $\mathbb{P}_N : L^2(I^d) \to S_N^d$, so that

$$\begin{aligned} \mathbb{P}_N u &= \sum_{|\vec{\mathbf{k}}|_1 \leq N} \hat{u}_{\vec{\mathbf{k}}}\mathbf{L}_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}) \\ \hat{u}_{\vec{\mathbf{k}}} &= \prod_{i=1}^{d}(k_i + \frac{1}{2})\int_{I^d} u(\vec{\mathbf{x}})\mathbf{L}_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})d\vec{\mathbf{x}} \end{aligned} \tag{14}$$

for all $u \in L^2(I^d)$, where $|\vec{\mathbf{k}}|_1 := \sum_{i=1}^d k_i$.

Throughout this article, we focus on the general bounded domain defined as $\Omega \subset \mathbb{R}^d$ and adopt the standard notations $W^{r,p}$ for Sobolev spaces with norm $||\cdot||_{W^{r,p}}$ and the seminorm $|\cdot|_{W^{r,p}}$ (see [1]). Specially, we set $W_0^{r,p}(\Omega) = \{u \in W^{r,p} : u|_{\partial\Omega} = 0\}$. If $p = 2$, we denote $W^{r,2}$ and $W_0^{r,2}$ as H^r and H_0^r, respectively. For any $r \in \mathbb{Z}_+$, we have

$$H^r := \left\{u :\ D^\alpha u \in L^2(\Omega), \forall \alpha \in \mathbb{Z}_+^d, 0 \leq |\alpha|_1 \leq r\right\}$$

where $\alpha = (\alpha_1,\ldots,\alpha_d)$, $\alpha_i \in \mathbb{Z}_+$, $i = 1,\ldots,d$, and $|\alpha|_1 = \sum_{i=1}^d \alpha_i$, and $D^\alpha u := \frac{\partial^\alpha u}{\partial x_1^{\alpha_1}\cdots\partial x_d^{\alpha_d}}$. It is clear that $H^0(\Omega) = L^2(\Omega)$. Actually, the Sobolev space $H^r(\Omega)$ is equipped with the seminorm and norm defined as

$$\begin{aligned} |u|_{H^\alpha(\Omega)}^2 &:= \langle D^\alpha u, D^\alpha u\rangle,\ 0 \leq |\alpha|_1 \leq r \\ ||u||_{H^r(\Omega)}^2 &:= \sum_{|\alpha|_1=0}^{r} |u|_{H^\alpha(\Omega)}^2 \end{aligned} \tag{15}$$

where $u \in H^r(\Omega)$. The index r can be viewed as the indicator of the regularity of the functions.

In addition, we consider the operators of orthogonal projection for the inner product of the Sololev spaces $H^1(I^d)$ and $H_0^1(I^d)$. Define the space $V_N^d := \{v \in S_N^d : v|_{\partial(I^d)} = 0\}$. Analogously,

define $\mathbb{P}_N^{1,0}: H_0^1(I^d) \to V_N^d$, such that $[\mathbb{P}_N^{1,0}u, \phi] = [u, \phi]$, for any $\phi \in V_N^d$, where $[u,v] = \int_{I^d} \nabla u \cdot \nabla v d\vec{\mathbf{x}}$ denotes the inner product of $H_0^1(I^d)$. In the bounded domain, this norm is equivalent to the usual $H_0^1(I^d)$ norm, due to the Poincaré inequality.

The following lemma is important for deriving a posterior error estimate. It can be found in the reference book [9].

Lemma 1: For all $u \in H^r(I^d), r \geq 0$, we have

$$||u - \mathbb{P}_N u||_{H^l(I^d)} \lesssim N^{\sigma(l)-r}||u||_{H^r(I^d)} \tag{16}$$

for $0 \leq l \leq r$, where

$$\sigma(l) = \begin{cases} 0, & \text{if } l = 0, \\ 2l - \frac{1}{2}, & \text{if } l > 0 \end{cases}$$

and $\lesssim$ represents less than or equal to with a generic constant c in front.

Moreover, if $u \in H^r(I^d) \bigcap H_0^1(I^d)$ with $r \geq 1$, set $U_N := \mathbb{P}_N^{1,0}u$, then the following estimates hold:

$$||u - U_N||_{H^\mu(I^d)} \lesssim N^{\mu-r}||u||_{H^r(I^d)} \tag{17}$$

for $\mu = 0, 1$.

It is clear to see that the projection onto V_N^d has a faster convergence rate than that onto S_N^d, if the function u vanishes on the boundary $\partial(I^d)$.

III. Legendre Galerkin Spectral Method in NLF Problems

Recall the brief description of on- and off-line algorithm in the introduction, the off-line computation is to numerically solve the FKE (9) on each interval $[\tau_i, \tau_{i+1}]$, $0 = \tau_0 < \tau_1 < \cdots < \tau_k = T$. Without loss of generality, we denote the bounded domain as $[-1,1]^d \times [0,\tau]$ after appropriate scaling in spatial domain Ω. Therefore, we shall use LSM to solve the following initial-boundary value problem (I-BVP):

$$\begin{cases} \frac{\partial u}{\partial t}(\vec{\mathbf{x}}, t) = \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j}\left[(GQG^\top)_{ij}u\right] - \sum_{i=1}^d \frac{\partial (f_i u)}{\partial x_i} \\ \qquad\qquad - \frac{1}{2}h^\top S^{-1} h u, \ (\vec{\mathbf{x}}, t) \in \Omega \times [0,\tau] \\ u(\vec{\mathbf{x}}, t)|_{\partial\Omega} = 0, \ t \in [0,\tau] \\ u(\vec{\mathbf{x}}, 0) = u_0(\vec{\mathbf{x}}), \ \vec{\mathbf{x}} \in \Omega. \end{cases} \tag{18}$$

Remark 1: If without any special explanations, the notation of $\vec{\mathbf{x}}$ in (18) and in the sequel has the same meaning as $x \in \mathbb{R}^d$ in Section II-A, which represents the state x_t of the signal observation system (1) after ignoring the subscript t.

The well-posedness of (18) is quite standard, since it is a linear parabolic PDE in bounded domain. In the remaining of this section, we shall investigate the convergence rate of the LSM in solving (18). The weak formulation is to find $u^N(\cdot,\cdot): \Omega \times [0,\tau) \to \mathbb{R}$, i.e., $u^N(\cdot,t) \in V_N^d$, for any $t \in [0,\tau]$ such that

$$\begin{cases} \left\langle \partial_t u^N, \phi \right\rangle = \frac{1}{2}\sum_{i,j=1}^d \left\langle \frac{\partial^2}{\partial x_i \partial x_j}\left[(GQG^\top)_{ij}u^N\right], \phi \right\rangle \\ \qquad\qquad - \sum_{i=1}^d \left\langle \frac{\partial (f_i u^N)}{\partial x_i}, \phi \right\rangle - \frac{1}{2}\left\langle h^\top S^{-1} h u^N, \phi \right\rangle \\ u^N(\vec{\mathbf{x}}, 0) = \mathbb{P}_N u_0(\vec{\mathbf{x}}) \end{cases} \tag{19}$$

for all $\phi \in V_N^d$. The Dirichlet boundary condition is satisfied by u^N naturally, due to the definition of V_N^d. In the following theorem, we shall investigate the convergence rate with respect to N.

Theorem 1: We assume that on $\Omega \times [0,\tau]$

1) $|\frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j}(GQG^\top)_{ij} - \sum_{i=1}^d \frac{\partial f_i}{\partial x_i}| \leq M_1$
2) $|h^\top S^{-1} h| \leq M_2$
3) $\|[\frac{1}{2}\sum_{j=1}^d \frac{\partial}{\partial x_j}(GQG^\top)_{ij} - f_i]_{i=1}^d\|_2 \leq M_3$
4) $||K||_\infty \leq \infty$ such that $KK^\top = GQG^\top$

where $||\circ||_2 = (\circ_1^2 + \cdots + \circ_d^2)^{\frac{1}{2}}$, and $||\circ||_\infty = \max_{ij}\{\circ_{ij}\}$. If $u \in L^2([0,\tau]; H^r(\Omega))$, for some $r > 1$, is the solution to (18), then

$$||u - u^N||^2(t) \leq c^* N^{1-r}, \ t \in [0,\tau] \tag{20}$$

where c^* depends only on $\tau, d, ||K||_\infty, M_i, i = 1,2,3$, and $\int_0^t ||u||^2_{H^r(\Omega)}(s)ds$.

Proof: Denote $U_N := \mathbb{P}_N^{0,1}u$ as in Lemma 1. For all $\phi \in V_N^d$, we have

$$\begin{aligned} 0 = \ & \langle \partial_t (u - U_N), \phi \rangle \\ \Rightarrow \quad \langle \partial_t U_N, \phi \rangle \overset{(3.18)}{=} & \frac{1}{2}\sum_{i,j=1}^d \left\langle \frac{\partial^2}{\partial x_i \partial x_j}\left[(GQG^\top)_{ij}u\right], \phi \right\rangle \\ & - \sum_{i=1}^d \left\langle \frac{\partial (f_i u)}{\partial x_i}, \phi \right\rangle \\ & - \frac{1}{2}\left\langle h^\top S^{-1} h u, \phi \right\rangle. \end{aligned} \tag{21}$$

Combine with (19) by linearity, it yields that

$$\begin{aligned} & \left\langle \partial_t (u^N - U_N), \phi \right\rangle \\ & = \frac{1}{2}\sum_{i,j=1}^d \left\langle \frac{\partial^2}{\partial x_i \partial x_j}\left[(GQG^\top)_{ij}(u^N - u)\right], \phi \right\rangle \\ & \quad - \sum_{i=1}^d \left\langle \frac{\partial [f_i (u^N - u)]}{\partial x_i}, \phi \right\rangle \\ & \quad - \frac{1}{2}\left\langle h^\top S^{-1} h (u^N - u), \phi \right\rangle \end{aligned} \tag{22}$$

for all $\phi \in V_N^d$. Let us denote $\varrho_N := u^N - U_N$, and choose $\phi = \varrho_N$, then by the integration by parts, we

have

$$\begin{aligned}
&\partial_t||\varrho_N||^2\\
&\overset{(3.22)}{=}\left\langle 2\varrho_N,\left[\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}(GQG^\top)_{ij}\right.\right.\\
&\qquad\left.\left.-\sum_{i=1}^{d}\frac{\partial f_i}{\partial x_i}-\frac{1}{2}h^\top S^{-1}h\right]\varrho_N\right\rangle\\
&+\sum_{i=1}^{d}\left\langle 2\varrho_N,\left[\frac{1}{2}\sum_{j=1}^{d}\frac{\partial}{\partial x_i}(GQG^\top)_{ij}-f_i\right]\frac{\partial\varrho_N}{\partial x_i}\right\rangle\\
&-\sum_{i=1}^{d}\left\langle\sum_{j=1}^{d}(GQG^\top)_{ij}\frac{\partial\varrho_N}{\partial x_j},\frac{\partial\varrho_N}{\partial x_i}\right\rangle\\
&+\left\langle 2\varrho_N,\left[\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}(GQG^\top)_{ij}\right.\right.\\
&\qquad\left.\left.-\sum_{i=1}^{d}\frac{\partial f_i}{\partial x_i}-\frac{1}{2}h^\top S^{-1}h\right](U_N-u)\right\rangle\\
&+\sum_{i=1}^{d}\left\langle 2\varrho_N\left[\frac{1}{2}\sum_{j=1}^{d}\frac{\partial}{\partial x_i}(GQG^\top)_{ij}-f_i\right]\frac{\partial(U_N-u)}{\partial x_i}\right\rangle\\
&-\sum_{i=1}^{d}\left\langle\sum_{j=1}^{d}(GQG^\top)_{ij}\frac{\partial(U_N-u)}{\partial x_j},\frac{\partial\varrho_N}{\partial x_i}\right\rangle\\
&:=I_1+I_2+I_3+I_4+I_5+I_6.
\end{aligned}\tag{23}$$

In the sequel, we estimate I_1-I_6 one-by-one

$$\begin{aligned}
I_1\le&\left|\left\langle 2\varrho_N,\left[\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}(GQG^\top)_{ij}\right.\right.\right.\\
&\left.\left.\left.-\sum_{i=1}^{d}\frac{\partial f_i}{\partial x_i}-\frac{1}{2}h^\top S^{-1}h\right]\varrho_N\right\rangle\right|\\
\le&\int_\Omega 2|\varrho_N|^2\left|\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}(GQG^\top)_{ij}\right.\\
&\left.-\sum_{i=1}^{d}\frac{\partial f_i}{\partial x_i}-\frac{1}{2}h^\top S^{-1}h\right|d\vec{\mathbf{x}}\\
\le&(2M_1+M_2)||\varrho_N||^2
\end{aligned}\tag{24}$$

by Assumptions 1) and 2). In I_2, it yields that

$$\begin{aligned}
I_2=&\sum_{i=1}^{d}\left\langle 2\varrho_N,\left[\frac{1}{2}\sum_{j=1}^{d}\frac{\partial}{\partial x_j}(GQG^\top)_{ij}-f_i\right]\frac{\partial\varrho_N}{\partial x_i}\right\rangle\\
=&\sum_{i=1}^{d}\int_\Omega\left[\frac{1}{2}\sum_{j=1}^{d}\frac{\partial}{\partial x_j}(GQG^\top)_{ij}-f_i\right]\frac{\partial\varrho_N^2}{\partial x_i}d\vec{\mathbf{x}}\\
=&-\int_\Omega\varrho_N^2\left[\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}(GQG^\top)_{ij}-\sum_{i=1}^{d}\frac{\partial f_i}{\partial x_i}\right]d\vec{\mathbf{x}}
\end{aligned}$$

where the last equality follows by integration by parts. By Assumption 1), it can be controlled by $||\varrho_N||$

$$\begin{aligned}
I_2\le&\int_\Omega\varrho_N^2\left|\frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i\partial x_j}(GQG^\top)_{ij}-\sum_{i=1}^{d}\frac{\partial f_i}{\partial x_i}\right|d\vec{\mathbf{x}}\\
\le&M_1||\varrho_N||^2.
\end{aligned}\tag{25}$$

In addition, we have

$$\begin{aligned}
I_3=&-\sum_{i=1}^{d}\left\langle\sum_{j=1}^{d}(GQG^\top)_{ij}\frac{\partial\varrho_N}{\partial x_j},\frac{\partial\varrho_N}{\partial x_i}\right\rangle\\
=&-\int_\Omega(\nabla\varrho_N)^\top GQG^\top\nabla\varrho_N d\vec{\mathbf{x}}\\
=&-\int_\Omega(K^\top\nabla\varrho_N)^\top(K^\top\nabla\varrho_N)d\vec{\mathbf{x}}=-||K^\top\nabla\varrho_N||^2.
\end{aligned}\tag{26}$$

With the similar argument as in I_1, we have

$$\begin{aligned}
I_4\le&\left(M_1+\frac{M_2}{2}\right)\int_\Omega 2|\varrho_N||U_N-u|d\vec{\mathbf{x}}\\
\le&\left(M_1+\frac{M_2}{2}\right)(||\varrho_N||^2+||U_N-u||^2)
\end{aligned}\tag{27}$$

by Assumptions 1), 2), and Cauchy–Schwartz inequality

$$\begin{aligned}
I_5\le&M_3\int_\Omega 2|\varrho_N||\nabla(U_N-u)|d\vec{\mathbf{x}}\\
\le&M_3(||\varrho_N||^2+||\nabla(U_N-u)||^2)
\end{aligned}\tag{28}$$

by Assumption 3) and Cauchy–Schwartz inequality. As for I_6, we proceed similarly

$$\begin{aligned}
I_6=&\left|\int_\Omega(\nabla\varrho_N)^\top(GQG^\top)\nabla(U_N-u)d\vec{\mathbf{x}}\right|\\
\le&||K^\top\nabla\varrho_N||^2+||K^\top\nabla(U_N-u)||^2\\
\le&||K^\top\nabla\varrho_N||^2+||K||_\infty^2||\nabla(U_N-u)||^2
\end{aligned}\tag{29}$$

by Cauchy–Schwartz inequality and Assumption 4). Substituting (24)–(29) back into (23), we have

$$\begin{aligned}
\partial_t||\varrho_N||^2\le&\left(4M_1+\frac{3M_2}{2}+M_3\right)||\varrho_N||^2\\
&+\left(M_1+\frac{M_2}{2}\right)||U_N-u||^2\\
&+\left(M_3+||K||_\infty^2\right)||\nabla(U_N-u)||^2.
\end{aligned}\tag{30}$$

Therefore, we obtain that

$$\begin{aligned}
&\partial_t||\varrho_N||^2-\left(4M_1+\frac{3M_2}{2}+M_3\right)||\varrho_N||^2\\
&\le\left(M_1+\frac{M_2}{2}\right)||U_N-u||^2\\
&\quad+\left(M_3+||K||_\infty^2\right)||\nabla(U_N-u)||^2\\
&\le\widehat{C}(||U_N-u||^2+||\nabla(U_N-u)||^2)=\widehat{C}||U_N-u||_{H^1(\Omega)}^2
\end{aligned}\tag{31}$$

where $\widehat{C} = \max\{M_1 + \frac{M_2}{2}, M_3 + ||K||_\infty^2\}$. By Gronwall's inequality, (31) yields that

$$\begin{aligned}&\frac{\partial}{\partial t}\left[e^{-\int_0^t(4M_1+\frac{3M_2}{2}+M_3)ds}||\varrho_N||^2(t)\right]\\&\quad\le \widehat{C}e^{-\int_0^t(4M_1+\frac{3M_2}{2}+M_3)ds}||U_N-u||^2_{H^1(\Omega)}(t)\\&\quad\le \widehat{C}||U_N-u||^2_{H^1(\Omega)}(t)\end{aligned}\tag{32}$$

where the last inequality is based on the fact that $e^{-\int_0^t(4M_1+\frac{3M_2}{2}+M_3)ds}\le 1$, for any $t\in[0,\tau]$. Integrating over $[0,t]$ on both sides of (32), we obtain

$$\begin{aligned}&e^{-\int_0^t(4M_1+\frac{3M_2}{2}+M_3)ds}||\varrho_N||^2(t)-||\varrho_N||^2(0)\\&\quad\le\int_0^t\widehat{C}||U_N-u||^2_{H^1(\Omega)}(s)ds.\end{aligned}\tag{33}$$

That is

$$\begin{aligned}||\varrho_N||^2(t)&\le e^{(4M_1+\frac{3M_2}{2}+M_3)t}\\&\quad\cdot\left(||\varrho_N||^2(0)+\int_0^t\widehat{C}||U_N-u||^2_{H^1(\Omega)}(s)ds\right)\\&\le\widetilde{C}\int_0^t||U_N-u||^2_{H^1(\Omega)}(s)ds\overset{(2.17)}{\le}CN^{1-r}\end{aligned}\tag{34}$$

where the generic constant C in front depends on τ, d, $||K||_\infty$, M_i, $i=1,2,3$, and the norm of $L^2([0,\tau];H^r(\Omega))$, but not N. By the triangular inequality and Lemma 1 again, we have

$$||u-u^N||^2(t)\le||\varrho_N||^2+||u-U_N||^2\le c^*N^{1-r}\tag{35}$$

where c^* is a constant depending on τ, d, $||K||_\infty$, M_is in Assumption 1)–3), and $\int_0^t||u||^2_{H^r(\Omega)}(s)ds$. ■

Remark 2: The convergence rate of LSM to FKE is twice faster than the HSM in [29], under the suitable conditions.

Remark 3: In fact, the Legendre Galerkin spectral method, proposed for NLF problems in this article, is defined in a compact domain, say $[-1,1]^d$ after scaling and translating. Thus, Assumption 1)–4) can be satisfied, say if f,h, and G are continuous. In bounded domain, the boundedness conditions are easier to be satisfied than those in unbounded one, due to the compactness.

IV. Numerics

A. Detailed Formulation

Lemma 2: (Lemma 2.1, [38]) Let us define the univariate generalized Legendre polynomials as

$$\begin{aligned}\psi_k(x)&:=a_k(L_k(x)-L_{k+2}(x))\\a_k&=\frac{1}{\sqrt{4k+6}}\end{aligned}\tag{36}$$

for $x\in[-1,1]$, L_ks are the Legendre polynomials defined in Section II. Let us denote

$$\begin{aligned}b_{j,k}&:=\langle\psi_k(x),\psi_j(x)\rangle\\c_{j,k}&:=\left\langle\psi_k'(x),\psi_j'(x)\right\rangle\\d_{j,k}&:=\left\langle\psi_k(x),\psi_j'(x)\right\rangle.\end{aligned}\tag{37}$$

Given an integer $N\ge2$, and for $0\le j,k\le N-2$, we have

$$b_{j,k}=b_{k,j}=\begin{cases}a_ka_j\left(\frac{2}{2j+1}+\frac{2}{2j+5}\right), & \text{if } k=j\\ a_ka_j\frac{2}{2k+1}, & \text{if } k=j+2\\ 0, & \text{otherwise}\end{cases}\tag{38}$$

$$c_{j,k}=\begin{cases}1, & \text{if } j=k\\ 0, & \text{otherwise}\end{cases}\tag{39}$$

and

$$d_{j,k}=-d_{k,j}=\begin{cases}-2a_ja_{j+1}, & \text{if } k=j+1\\ 2a_{j-1}a_j, & \text{if } k=j-1\\ 0, & \text{otherwise.}\end{cases}\tag{40}$$

Consequently, the linear space spanned by the multivariate generalized Legendre polynomials is

$$\begin{aligned}V_N^d=\text{span}\{&\Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})=\psi_{k_1}(x_1)\psi_{k_2}(x_2)\cdots\psi_{k_d}(x_d):\\&\vec{\mathbf{x}}=(x_1,\ldots,x_d)\in I^d,\vec{\mathbf{k}}=(k_1,\ldots,k_d)\\&0\le k_j\le N-2,j=1,2,\ldots,d\}\end{aligned}\tag{41}$$

which is the one defined in Section II.

It is clear that (9) is independent of the observation path $\{y_{\tau_i}\}_{i=1}^k$, and the transformation between u_i and ρ_i is one-to-one. It is also not hard to see that (9) could be numerically solved beforehand. Let us denote the second elliptic operator in (9) as $A(t)$ for short. Under some conditions, $\{A(t)\}_{t\in[0,T]}$ forms a family of strong elliptic operators. Furthermore, the operator $A(t):D(A(t))\subset L^2(\mathbb{R}^d)\longrightarrow L^2(\mathbb{R}^d)$ is the infinitesimal generator of the two-parameter semigroup $\mathcal{A}(t,\tau)$, for $t>\tau$. In particular, we obtain a sequence of two-parameter semigroup $\{\mathcal{A}(t,\tau_{i-1})\}_{i=1}^k$, for $\tau_{i-1}\le t<\tau_i$.

However, as we motivated in Section I, the state is generally defined in some bounded domain $\Omega\in\mathbb{R}^d$. Thus, we shall solve the I-BVP problem (18) in each time interval $[\tau_{i-1},\tau_i]$ after proper scaling in spatial domain Ω, which is essentially the off-line part of our algorithm.

The off-line part of computation is first to take the initial conditions of (18) at $t=\tau_i$ as a set of complete orthonormal bases in $L^2(I^d)$ with vanishing boundary condition, where I^d is the support of the orthonormal bases in L^2, and then to seek the approximate weak solution of (18) in V_N^d expressed by

$$u^N(\vec{\mathbf{x}},t)\approx\sum_{0\le|\vec{\mathbf{k}}|_1\le N-2}q_{\vec{\mathbf{k}}}(t)\Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}),t\in[\tau_{i-1},\tau_i)$$

where $\Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})\in V_N^d$ are the generalized Legendre polynomials defined in Lemma 2, $q_{\vec{\mathbf{k}}}(t)\in\mathbb{R}^{N-1}$ are the coefficients, and

$|\vec{\mathbf{k}}|_1 = \sum_{j=1}^{d} k_j$, k_js are integers with $0 \le k_j \le N-2$. For the simplicity of notation, we ignore the superscript of $u^N(\vec{\mathbf{x}}, t)$ as $u(\vec{\mathbf{x}}, t)$ in the sequel.

Remark 4: The construction of the polynomial bases with vanishing boundary condition from Legendre polynomials is not unique. However, the ones constructed in Lemma 2 are easy to implement.

The on-line computation in our algorithm is consisted of two parts at each time step $\tau_{i-1}, i = 1, \ldots, k$.

1) Project the initial condition $u_i(\vec{\mathbf{x}}, \tau_{i-1}) \in L^2(I^d)$ at $t = \tau_{i-1}$ onto the bases $\{\Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})\}_{0 \le |\vec{\mathbf{k}}|_1 \le N-2}$, i.e.,

$$u_i(\vec{\mathbf{x}}, \tau_{i-1}) = \sum_{0 \le |\vec{\mathbf{k}}|_1 \le N-2} q^i_{\vec{\mathbf{k}}} \Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}}).$$

Hence, the solution to (9) at $t = \tau_i$ can be expressed as

$$\begin{aligned} u_i(\vec{\mathbf{x}}, \tau_i) &= \mathcal{A}(\tau_i, \tau_{i-1}) u_i(\vec{\mathbf{x}}, \tau_{i-1}) \\ &= \sum_{0 \le |\vec{\mathbf{k}}|_1 \le N-2} q^i_{\vec{\mathbf{k}}} \left[\mathcal{A}(\tau_i, \tau_{i-1}) \Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})\right] \end{aligned} \tag{42}$$

where $\{\mathcal{A}(\tau_i, \tau_{i-1}) \Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})\}_{0 \le |\vec{\mathbf{k}}|_1 \le N-2}$ have been calculated off-line.

2) Update the initial condition of (9) at τ_i with new observation y_{τ_i}. Let us specify the observation updates for each time step. For $0 \le t \le \tau_1$, the initial condition is $u_1(\vec{\mathbf{x}}, 0) = \sigma_0(\vec{\mathbf{x}})$. At time $t = \tau_1$, when the observation y_{τ_1} is available

$$\begin{aligned} & u_2(\vec{\mathbf{x}}, \tau_1) \\ & \overset{(2.8)}{=} \exp[h^\top(\vec{\mathbf{x}}, \tau_1) S^{-1}(\tau_1) y_{\tau_1}] \rho_2(\vec{\mathbf{x}}, \tau_1) \\ & \overset{(2.8),(2.6)}{=} \exp[h^\top(\vec{\mathbf{x}}, \tau_1) S^{-1}(\tau_1) y_{\tau_1}] u_1(\vec{\mathbf{x}}, \tau_1) \end{aligned} \tag{43}$$

with the fact that $y_0 = 0$. Here

$$u_1(\vec{\mathbf{x}}, \tau_1) = \sum_{0 \le |\vec{\mathbf{k}}|_1 \le N-2} q^1_{\vec{\mathbf{k}}} \left[\mathcal{A}(\tau_1, 0) \Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})\right]$$

where $\{q^1_{\vec{\mathbf{k}}}\}_{0 \le |\vec{\mathbf{k}}|_1 \le N-2}$ is computed in the previous time step, and $\{\mathcal{A}(\tau_1, 0) \Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})\}_{0 \le |\vec{\mathbf{k}}|_1 \le N-2}$ are prepared by off-line computations. Hence, we obtain the initial condition $u_2(\vec{\mathbf{x}}, \tau_1)$ of (9) for the next time interval $\tau_1 \le t \le \tau_2$. Recursively, the initial condition of (9) for $\tau_{i-1} \le t < \tau_i$ is

$$\begin{aligned} & u_i(\vec{\mathbf{x}}, \tau_{i-1}) \\ & \quad = \exp[h^\top(\vec{\mathbf{x}}, \tau_{i-1}) S^{-1}(\tau_{i-1})(y_{\tau_{i-1}} - y_{\tau_{i-2}})] \\ & \qquad \cdot u_{i-1}(\vec{\mathbf{x}}, \tau_{i-1}) \end{aligned} \tag{44}$$

for $i = 2, 3, \ldots, k$, where $u_{i-1}(\vec{\mathbf{x}}, \tau_{i-1}) = \sum_{0 \le |\vec{\mathbf{k}}|_1 \le N-2} q^{i-2}_{\vec{\mathbf{k}}} [\mathcal{A}(\tau_{i-1}, \tau_{i-2}) \Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})]$.

At each time step $\tau_{i-1}, i = 1, \ldots, k$, as a PDF, the solution of FKE must satisfy the following constraint:

$$\lim_{\vec{\mathbf{x}} \to \infty} u_i(\vec{\mathbf{x}}, \tau_{i-1}) = 0, \text{ for } i = 1, \ldots, k \tag{45}$$

(45) is called the vanishing boundary condition. Any discretization-based method requires a compact domain for implementation, due to which the vanishing boundary condition is imposed on a conservatively chosen large enough domain.

Remark 5: For any PDF $p(\vec{\mathbf{x}}, t)$, the normality constraint denoted by $\int_{\mathbb{R}^d} p(\vec{\mathbf{x}}, t) = 1$ must hold at any time t. In the implementation of our algorithm to NLF problems, however, the normality condition at each temporal step can not be maintained, since the solution of FKE is numerically approximated by the LSM with the states at a large enough but bounded domain. As shown in the Algorithms 1–3, what is adopted to alleviate the normality constraint is to normalize the approximate solution of FKE at each time instance, especially see step 3 in Algorithm 1.

Remark 6: Compared with the time-invariant case, the price to pay is that the time-varying case requires more storage capacity. Since $\{\mathcal{A}(\tau_{i+1}, \tau_i) \Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})\}_{0 \le |\vec{\mathbf{k}}|_1 \le N-2}$ differs at each τ_i, $i = 1, \ldots, k$, and all of them need to be stored. In general, the longer simulation time is, the more storage it requires in the time-varying case. While the storage of the data is independent of the simulation time in the time-invariant case. Nevertheless, it would not affect the off-line virtue of our algorithm.

B. Numerical Experiments

In this section, we shall apply our algorithm to both time-invariant and time-varying case. The numerical simulations support our main theorem. In our implementation, we adopt the LSM with $\{\Psi_{\vec{\mathbf{k}}}\}$ to get the approximate solution of (9) in V_N^d in Lemma 2.

To compare the performance of different methods, we introduce the mean of the squared estimation error (MSE), and the MSE for one realization is defined as follows:

$$MSE_{x_l} = \frac{1}{k+1} \sum_{i=0}^{k} (x_l(\tau_i) - \hat{x}_l(\tau_i))^2 \tag{46}$$

where $x_l(\tau_i)$ is the lth component of the real state $\vec{\mathbf{x}}$ at the instant τ_i and $\hat{x}_l(\tau_i)$ is the estimation of $x_l(\tau_i)$, for $l = 1, \ldots, d$.

The off-line part of our algorithm is to numerically solve the FKE, i.e., to obtain the weak solution of FKE by generalized Legendre polynomials. By taking the approximation to the weak solution of (19) as

$$u^N(\vec{\mathbf{x}}, t) \approx \sum_{0 \le |\vec{\mathbf{k}}|_1 \le N-2} a_{\vec{\mathbf{k}}}(t) \Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})$$

where $a_{\vec{\mathbf{k}}}(t) \in \mathbb{R}^{N-1}, \vec{\mathbf{k}} = (k_1, \ldots, k_d)$ are the coefficients. Let us denote

$$\Phi(x) := [\psi_0(x), \psi_1(x), \ldots, \psi_{N-2}(x)]$$

and

$$D\Phi(x) := \left[\frac{d\psi_0(x)}{dx}, \frac{d\psi_1(x)}{dx}, \ldots, \frac{d\psi_{N-2}(x)}{dx}\right]$$

where $\psi_i(x)$ is the univariate basis in $L(\mathbb{R}^1)$. Correspondingly, we have

$$u^N(\vec{\mathbf{x}}, t) \approx \left(\otimes_{i=1}^{d} \Phi(x_i)\right) a(t)$$

where $a(t) \in \mathbb{R}^{(N-1)^d}$, $\otimes_{i=1}^{d} \Phi(x_i) := \Phi(x_1) \otimes \cdots \otimes \Phi(x_d)$, and $\otimes$ is Kronecker product. The notations are defined as

following:

$$[\otimes_{k=1}^{d}\Phi(x_k)]_i := \otimes_{k=1}^{i-1}\Phi(x_k)\otimes D\Phi(x_i)\otimes_{k=i+1}^{d}\Phi(x_k) \tag{47}$$

$$\begin{aligned}&\left[\otimes_{k=1}^{d}\left(\Phi(x_k)^\top\Phi(x_k)\right)\right]_j\\ &:= \otimes_{k=1}^{j-1}(\Phi(x_k)^\top\Phi(x_k))\otimes(\Phi(x_j)^\top\Phi(x_j))\\ &\quad\otimes_{k=j+1}^{d}\left(\Phi(x_k)^\top\Phi(x_k)\right)\end{aligned} \tag{48}$$

and

$$\begin{aligned}&\left[\otimes_{k=1}^{d}\left(\Phi(x_k)^\top\Phi(x_k)\right)\right]_{ij}\\ &:= \begin{cases}\otimes_{k=1}^{i-1}\left(\Phi(x_k)^\top\Phi(x_k)\right)\otimes\left(D\Phi(x_i)^\top\Phi(x_i)\right)\\ \quad\otimes_{k=i+1}^{j-1}\left(\Phi(x_k)^\top\Phi(x_k)\right)\otimes\left(\Phi(x_j)^\top\Phi(x_j)\right)\\ \quad\otimes_{k=j+1}^{d}\left(\Phi(x_k)^\top\Phi(x_k)\right), & i<j\\ \otimes_{k=1}^{i-1}\left(\Phi(x_k)^\top\Phi(x_k)\right)\otimes\left(D\Phi(x_i)^\top D\Phi(x_i)\right)\\ \quad\otimes_{k=i+1}^{d}\left(\Phi(x_k)^\top\Phi(x_k)\right), & i=j\\ \otimes_{k=1}^{j-1}\left(\Phi(x_k)^\top\Phi(x_k)\right)\otimes\left(D\Phi(x_j)^\top\Phi(x_j)\right)\\ \quad\otimes_{k=j+1}^{i-1}\left(\Phi(x_k)^\top\Phi(x_k)\right)\otimes\left(\Phi(x_i)^\top\Phi(x_i)\right)\\ \quad\otimes_{k=i+1}^{d}\left(\Phi(x_k)^\top\Phi(x_k)\right), & i>j\end{cases}\end{aligned} \tag{49}$$

for any $i,j=1,\ldots,d$, respectively.

Recall that at each temporal interval, by taking $\phi(\vec{\mathbf{x}})=\Psi_{\vec{\mathbf{k}}}(\vec{\mathbf{x}})$ the weak solution of FKE (19) satisfies

$$\begin{cases}\left\langle\otimes_{k=1}^{d}\Phi(x_k),\otimes_{k=1}^{d}\Phi(x_k)^\top\right\rangle\dfrac{d}{dt}a(t)\\ \quad=-\dfrac{1}{2}\sum_{i,j=1}^{d}\left\langle(GQG^\top)_{ij}[\otimes_{k=1}^{d}\Phi(x_k)]_i,\right.\\ \quad\left.([\otimes_{k=1}^{d}\Phi(x_k)]_j)^\top\right\rangle a(t)\\ \quad-\dfrac{1}{2}\sum_{i,j=1}^{d}\left\langle\dfrac{\partial(GQG^\top)_{ij}}{\partial x_i}\otimes_{k=1}^{d}\Phi(x_k),\right.\\ \quad\left.([\otimes_{k=1}^{d}\Phi(x_k)]_j)^\top\right\rangle a(t)\\ \quad+\sum_{i=1}^{d}\left\langle f_i\otimes_{k=1}^{d}\Phi(x_k),([\otimes_{k=1}^{d}\Phi(x_k)]_i)^\top\right\rangle a(t)\\ \quad-\dfrac{1}{2}\left\langle h^\top S^{-1}h\otimes_{k=1}^{d}\Phi(x_k),\otimes_{k=1}^{d}\Phi(x_k)^\top\right\rangle a(t)\\ \quad:=(II_1+II_2+II_3+II_4)a(t)\\ u^N(\vec{\mathbf{x}},0)=\otimes_{k=1}^{d}\Phi(x_k)a(0)\end{cases} \tag{50}$$

where $a(0)\in\mathbb{R}^{(N-1)^d}$ is obtained from $u_0(\vec{\mathbf{x}})\approx\otimes_{i=1}^{d}\Phi(x_i)a(0)$.

In order to speed up our Galerkin spectral method to FKE by using generalized Legendre polynomials shown in Lemma 2, it is necessary to scale the general bounded domain Ω into I^d, i.e., for any rectangular domain $\vec{\mathbf{x}}=(x_1,\ldots,x_d)\in\Omega:=[M_1,N_1]\times\cdots\times[M_d,N_d]\subset\mathbb{R}^d$, taking $\widetilde{x}_i=\frac{2}{N_i-M_i}x_i+\frac{M_i+N_i}{M_i-N_i}$, $i=1,\ldots,d$, then it is obtained that $\widetilde{\mathbf{x}}=(\widetilde{x}_1,\ldots,\widetilde{x}_d)\in I^d=[-1,1]^d$. Correspondingly, we set $\widetilde{\Phi}(\tilde{x}):=\Phi(\tilde{x})$, and $D\widetilde{\Phi}(\tilde{x}):=D\Phi(\tilde{x})$, for any $\tilde{x}\in[-1,1]$.

Therefore, (50) leads to an ordinary differential equation (ODE) of $a(t)$ as follows:

$$\frac{\prod_{k=1}^{d}(N_k-M_k)}{2^d}\left(\otimes_{i=1}^{d}B\right)\frac{da(t)}{dt}=R(t)a(t),t\in[0,\tau]. \tag{51}$$

The left side of (51) is obtained from

$$\begin{aligned}&\left\langle\otimes_{i=1}^{d}\Phi(x_i),\otimes_{i=1}^{d}\Phi(x_i)^\top\right\rangle\\ &=\frac{\prod_{k=1}^{d}(N_k-M_k)}{2^d}\int_{I^d}\otimes_{i=1}^{d}\left(\widetilde{\Phi}(\widetilde{x}_i)^\top\widetilde{\Phi}(\widetilde{x}_i)\right)d\widetilde{\mathbf{x}}\\ &=\frac{\prod_{k=1}^{d}(N_k-M_k)}{2^d}\otimes_{i=1}^{d}B\end{aligned} \tag{52}$$

where the last equality is based on the fact of Lemma 2, and $B=(b_{kj})$ is defined in (38).

$R(t)$ in the right side of (51) consists of the following several parts:

$$\begin{aligned}II_1=&-\frac{1}{2}\frac{\prod_{k=1}^{d}(N_k-M_k)}{2^d}\sum_{i,j=1}^{d}\frac{2}{N_i-M_i}\frac{2}{N_j-M_j}\\ &\left\{\int_{I^d}(\widetilde{G}Q\widetilde{G}^\top)_{ij}\left[\otimes_{k=1}^{d}\left(\widetilde{\Phi}(\widetilde{x}_k)^\top\widetilde{\Phi}(\widetilde{x}_k)\right)\right]_{ij}d\widetilde{\mathbf{x}}\right\}\end{aligned} \tag{53}$$

$$\begin{aligned}II_2=&-\frac{1}{2}\frac{\prod_{k=1}^{d}(N_k-M_k)}{2^d}\sum_{i,j=1}^{d}\frac{2}{N_i-M_i}\frac{2}{N_j-M_j}\\ &\left\{\int_{I^d}\frac{\partial(\widetilde{G}Q\widetilde{G}^\top)_{ij}}{\partial\widetilde{x}_i}\left[\otimes_{k=1}^{d}\left(\widetilde{\Phi}(\widetilde{x}_k)^\top\widetilde{\Phi}(\widetilde{x}_k)\right)\right]_{j}d\widetilde{\mathbf{x}}\right\}\end{aligned} \tag{54}$$

$$\begin{aligned}II_3=&\frac{\prod_{k=1}^{d}(N_k-M_k)}{2^d}\sum_{i=1}^{d}\frac{2}{N_i-M_i}\\ &\times\left\{\int_{I^d}\widetilde{f}_i(\widetilde{\mathbf{x}})\left[\otimes_{k=1}^{d}\left(\widetilde{\Phi}(\widetilde{x}_k)^\top\widetilde{\Phi}(\widetilde{x}_k)\right)\right]_{i}d\widetilde{\mathbf{x}}\right\}\end{aligned} \tag{55}$$

and

$$\begin{aligned}II_4=&-\frac{1}{2}\frac{\prod_{k=1}^{d}(N_k-M_k)}{2^d}\\ &\times\left[\int_{I^d}\widetilde{h}^\top S^{-1}\widetilde{h}\otimes_{k=1}^{d}\left(\widetilde{\Phi}(\widetilde{x}_k)^\top\widetilde{\Phi}(\widetilde{x}_k)\right)d\widetilde{\mathbf{x}}\right].\end{aligned} \tag{56}$$

Thus, we obtain that

$$R(t):=\sum_{i=1}^{4}II_i. \tag{57}$$

Remark 7: (51) is an ODE about $a(t)$, and $t\in[0,\tau]$. Actually, it can be defined in any temporal interval $[s,l]$, for any $0\le s\le l$, and $s,l\in\mathbb{R}$. Correspondingly, $R(t)$ in (57) can be

Algorithm 3: Off-Line Part of LSM in 2-D Signal System.

1: **Initialization:** given $u_1(\vec{\mathbf{x}}, \tau_0)$ in Algorithm 1, then projecting $u_1(\vec{\mathbf{x}}, \tau_0)$ into V_N^2 as $u_1(\vec{\mathbf{x}}, \tau_0) = (\Phi_1 \otimes \Phi_2)^T a(\tau_0)$, we obtain $a(\tau_0)$.
2: **for** $i = 1$ to k **do**
3: Calculate $R(\tau_{i-1})$ in (57).
4: Obtain $a(\tau_{i-1})$ by updating the initial condition of (9) at τ_i with new arriving observation y_{τ_i} shown in Algorithm 1. Specifically, projecting $u_i(\vec{\mathbf{x}}, \tau_{i-1})$ into V_N^2 by $u_i(\vec{\mathbf{x}}, \tau_{i-1}) = (\Phi_1 \otimes \Phi_2)^T a(\tau_{i-1})$, we obtain $a(\tau_{i-1})$.
5: Calculate $\frac{N_1-M_1}{2}\frac{N_2-M_2}{2} B \otimes B \frac{da(t)}{dt} = R(\tau_{i-1})a(\tau_{i-1})$, $t \in [\tau_{i-1}, \tau_i]$.
6: Obtain $a(\tau_i)$ by letting $t = \tau_i$.
7: Obtain $u_i(\vec{\mathbf{x}}, \tau_i) = (\Phi_1 \otimes \Phi_2)^T a(\tau_i)$.
8: Normalize $u_i(\vec{\mathbf{x}}, \tau_i) = \frac{u_i(\vec{\mathbf{x}}, \tau_i)}{\int_\Omega u_i(\vec{\mathbf{x}}, \tau_i) d\mathbf{x}}$.
9: Obtain $\rho_i(\vec{\mathbf{x}}, \tau_i)$ by (8) and the normalized $u_i(\vec{\mathbf{x}}, \tau_i)$.
10: Normalize $\rho_i(\vec{\mathbf{x}}, \tau_i) = \frac{\rho_i(\vec{\mathbf{x}}, \tau_i)}{\int_\Omega \rho_i(\vec{\mathbf{x}}, \tau_i) d\vec{\mathbf{x}}}$.
11: **end for**

solved for any $t \in [\tau_{i-1}, \tau_i]$, $i = 1, 2, \ldots k$ in our algorithm. Actually, we use Gaussian quadratures to numerically compute II_i, $i = 1, \ldots 4$.

Remark 8: The off-line computation in our algorithm is to numerically solve the FKE (9), recursively in each time interval with different initial conditions. In this case, at each time interval $[\tau_{i-1}, \tau_i], i = 1, \ldots, k$, if the solution of FKE was projected into space V_N^d, i.e., the vector of unknowns $a(t)$ at each time step can be naturally considered as a d-dimensional array, like as $A(i_1, \ldots i_d), 0 \le i_l \le n_l, 1 \le l \le d$. For simplicity, we assume that all mode sizes n_k are equal to $N - 2$. The formal number of unknowns in this case behaves as $(N-1)^d$ and is subject to the curse of dimensionality. In the following section, when we utilized the Galerkin spectral method with generalized Legendre polynomials to numerically solve FKE, the curse of dimensionality also exists in the implementation of our algorithm.

C. Numerical Simulation for 2-D NLF Problems

We shall take $d = 2$ to illustrate the feasibility of our algorithm. By comparing with both EKF and PF methods, the experimental results show that our algorithm can achieve the real-time performance and good accuracy. We summarize the off-line part for 2-D NLF problem in Algorithm 3 below.

1) Time-Invariant System: Let us consider the 2-D cubic sensor problem in the spacial region of $[-1.4, 1.4] \times [-1.4, 1.4]$

$$\begin{cases} dx_1(t) = [-0.4x_1(t) + 0.1x_2(t)]\,dt + dv_1(t) \\ dx_2(t) = -0.6x_2(t)dt + dv_2(t) \\ dy_1(t) = x_1^3(t)dt + dw_1(t) \\ dy_2(t) = x_2^3(t)dt + dw_2(t) \end{cases} \tag{58}$$

TABLE I
PERFORMANCE OF DIFFERENT METHODS IN CUBIC SENSOR PROBLEMS

Methods	CPU time	MSE of state x_1	MSE of state x_2
Our algorithm with $N = 15^2$	1.43	0.4041	0.5402
Our algorithm with $N = 20^2$	3.86	0.4093	0.4244
EKF	0.32	$--$	$--$
PF	108.72	1.3907	0.9240

N means the total number of the elements in the set $\{\psi_i(x_1)\psi_j(x_2) : 0 \le i, j \le \sqrt{N} - 1\}$, i.e., the set $V^2_{\sqrt{N}-1}$.

where $x(t) = [x_1(t), x_2(t)]^\top \in \mathbb{R}^2$, $y(t) = [y_1(t), y_2(t)]^\top \in \mathbb{R}^2$, and $v(t) = [v_1(t), v_2(t)]^\top$, $w(t) = [w_1(t), w_2(t)]^\top$ are both 2-D Brownian motion processes with $E[dv(t)dv(t)^\top] = I_2 dt$, $E[dw(t)dw(t)^\top] = I_2 dt$, where I_2 is the 2-D identity matrix. In this simulation, the total experimental time is $T = 50$, and the time discretization $dt = 0.01$. The initial state x_0 is sampled from $\mathcal{N}((0.1, 0.12), 0.1I_2)$.

The corresponding FKE (9) in this case is

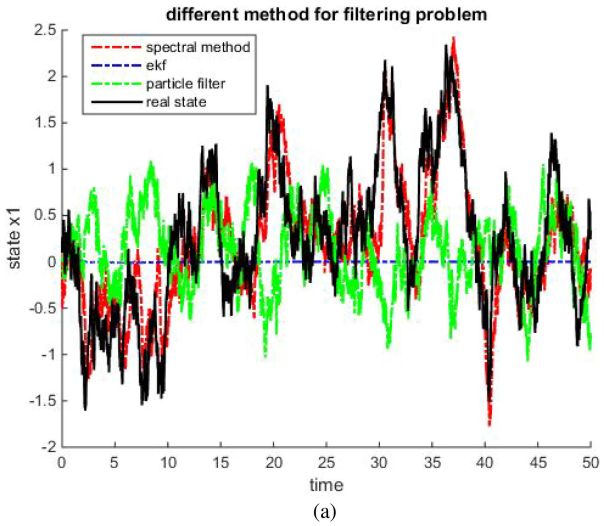
$$\begin{aligned} \frac{\partial u}{\partial t} &= \left(L - \frac{1}{2}h^\top S^{-1} h\right) u \\ &= \frac{1}{2}\left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}\right) + u - (-0.4x_1 + 0.1x_2)\frac{\partial u}{\partial x_1} \\ &\quad - (-0.6x_2)\frac{\partial u}{\partial x_2} - \frac{1}{2}\left[x_1(t)^6 + x_2(t)^6\right] u. \end{aligned} \tag{59}$$

Furthermore, we assume the initial distribution is $u_0(x) = \exp(-\frac{1}{2}|x|^2)$. This assumption is not crucial at all. The non-Gaussian ones, for instance $u_0(x) = \exp(-\frac{1}{2}|x|^4)$, gives the similar results. Meanwhile, the updated initial data are

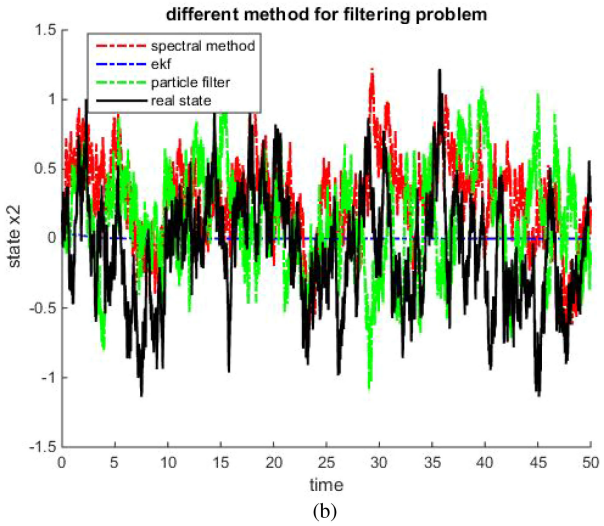
$$u_i(x, \tau_i) = \exp\left\{(x_1^3, x_2^3)^\top dy(t)\right\} u_{i-1}(x, \tau_i)$$

where $i = 1, \ldots, N_t$, and $N_t := T/dt = 5000$.

We compare our algorithm with EKF and PF with 100 particles. As shown in Figs. 1 and 2, which are the one realization of the states x_1, x_2, respectively, of the 20 Monte Carlo simulations, we see that our algorithm tracks the states better than PF, while the EKF completely fails on the experiments. The CPU times and the averaged MSE over 20 Monte Carlo simulations for different methods are displayed in Table I. When the total number of the generalized Legendre polynomials is taken to be $N = 15^2$, the on-line CPU time of our algorithm is only 1.43 s; consequently, the update time is less than 3×10^{-4}. Whereas the number is increased to $N = 20^2$, it costs our algorithm about 3.86 s to complete the simulation, i.e., the update time is around 7×10^{-4}. In conclusion, the LSM can be performed in real time, on account of the observation data coming in every 0.01s. The PF, which based on the Monte Carlo method to approximate the conditional distribution of the states, is much more time consuming and the MSE of both states are not satisfactory. EKF cannot even yield a reasonable estimation, due to the nonlinearity of the system. Theoretically, as the number of the generalized Legendre polynomials increases, the approximation error of u^N to u would decrease, thus the MSE of the states in our algorithm should decrease. However, as we can see that the MSE of x_1 by our algorithm with $N = 20^2$ is 5.2% higher than that with

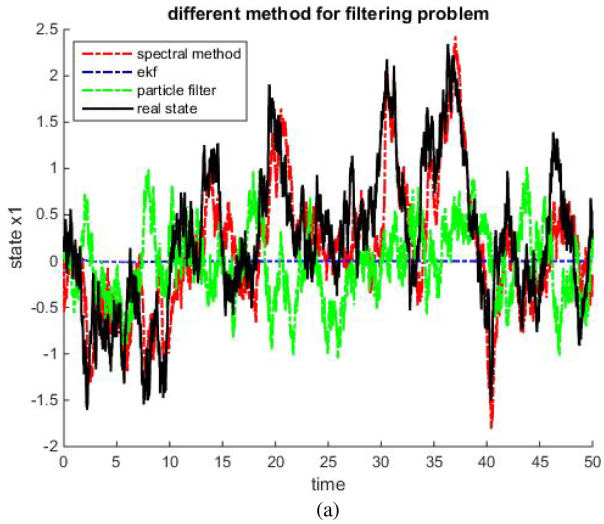


(a)

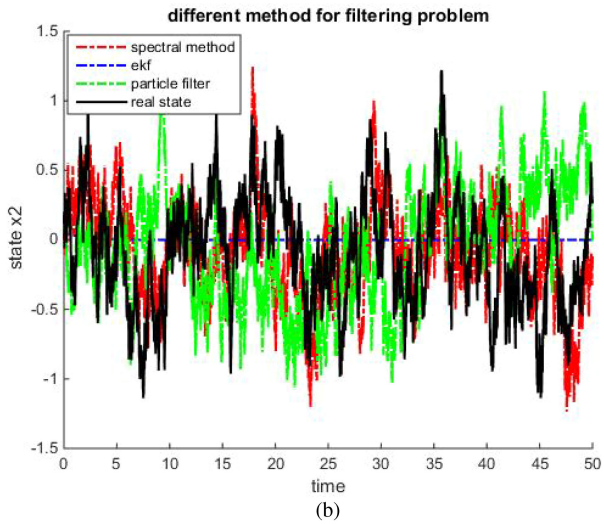


(b)

Fig. 1. State estimation results of different methods in cubic sensor with $N = 15^2$. (a) State x_1 with $N = 15^2$. (b) State x_2 with $N = 15^2$.



(a)



(b)

Fig. 2. State estimation results of different methods in cubic sensor with $N = 20^2$. (a) State x_1 with $N = 20^2$. (b) State x_2 with $N = 20^2$.

$N = 15^2$. This result indicates that the proper, not larger number of the generalized Legendre polynomials should be chosen in the implementation of our algorithm.

2) Time-Varying System: We consider the following time-varying almost linear sensor problem:

$$\begin{cases} dx_1(t) = [-0.4x_1(t) + 0.1x_2(t)]\,dt \\ \qquad\qquad + (1 + 0.1\cos(20\pi t))dv_1(t) \\ dx_2(t) = -0.6x_2(t)dt + [0.9 + 0.2\cos(18\pi t)]dv_2(t) \\ dy_1(t) = x_1(t)\,[1 + 0.2\cos(x_2(t))]\,dt + dw_1(t) \\ dy_2(t) = x_2(t)\,[1 + 0.2\cos(x_1(t))]\,dt + dw_2(t) \end{cases} \tag{60}$$

where $x(t) = [x_1(t), x_2(t)]^\top, y(t) = [y_1(t), y_2(t)]^\top \in \mathbb{R}^2$, and $v(t) = [v_1(t), v_2(t)]^\top, w(t) = [w_1(t), w_2(t)]^\top$ are both 2-D Brownian motion processes with $E[dv(t)dv(t)^\top] = Q(t)dt$, $E[dw(t)dw(t)^\top] = S(t)dt$. Let $Q(t)$ and $S(t)$ are the identity matrices in this simulation, and assume the initial state is sampled by the Gaussian distribution of $\mathcal{N}((0.1, 0.12), 0.1I_2)$.

The corresponding FKE (9) in this time-varying case is

$$\begin{aligned} \frac{\partial u}{\partial t} &= \left(L - \frac{1}{2}h^\top S^{-1}h\right)u \\ &= \frac{1}{2}[1 + 0.1\cos(20\pi t)]^2\frac{\partial^2 u}{\partial x_1^2} \\ &\quad + \frac{1}{2}[0.9 + 0.2\cos(18\pi t)]^2\frac{\partial^2 u}{\partial x_2^2} \\ &\quad + u - [-0.4x_1 + 0.1x_2]\frac{\partial u}{\partial x_1} - (-0.6x_2)\frac{\partial u}{\partial x_2} \\ &\quad - \frac{1}{2}x_1(t)^2[1 + 0.2\cos(x_2(t))]^2 u \\ &\quad - \frac{1}{2}x_2(t)^2[1 + 0.2\cos(x_1(t))]^2 u. \end{aligned}$$

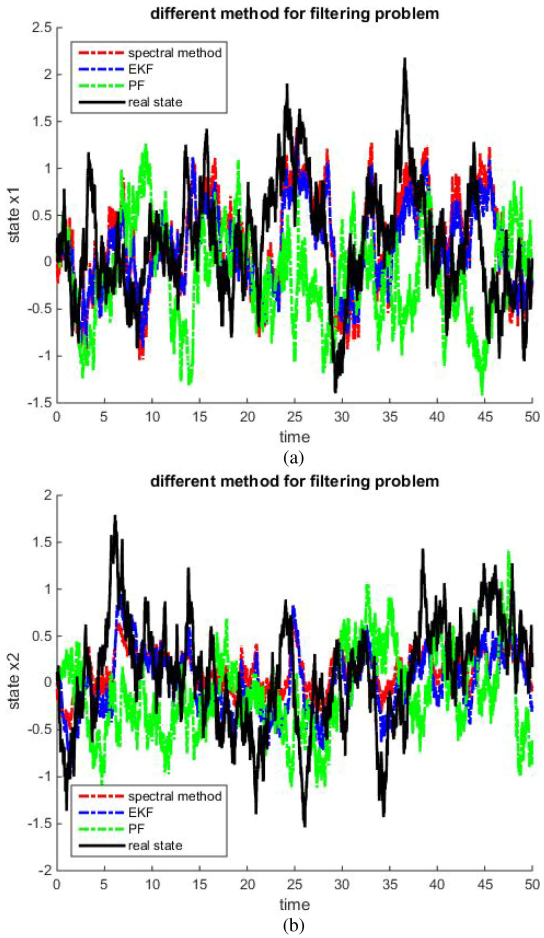


Fig. 3. State estimation results of different methods in time-varying almost linear sensor with $N = 5^2$. (a) State x_1 with $N = 5^2$. (b) State x_2 with $N = 5^2$.

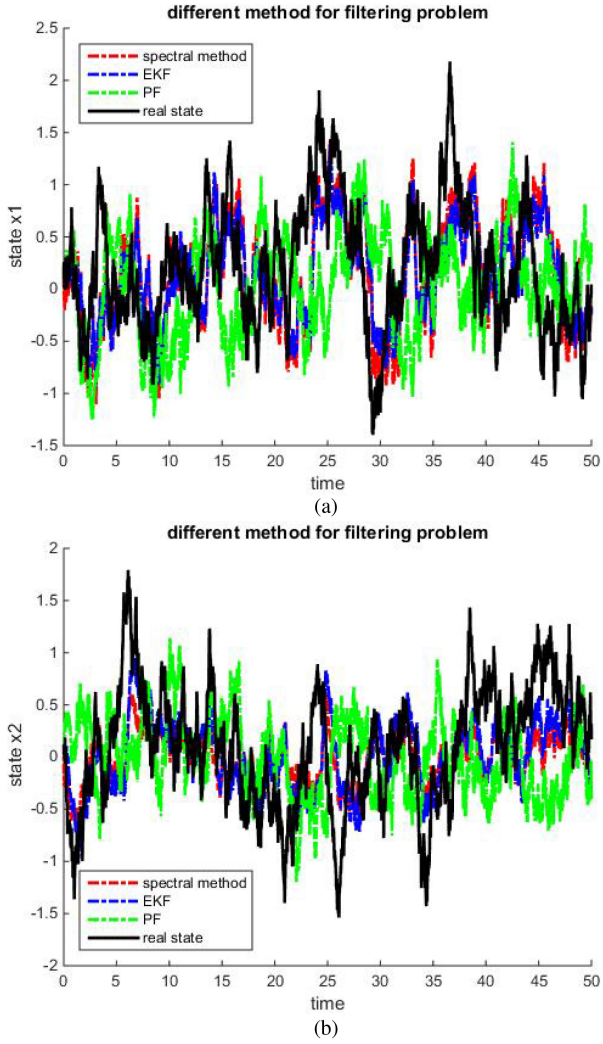


Fig. 4. State estimation results of different methods in time-varying almost linear sensor with $N = 10^2$. (a) State x_1 with $N = 10^2$. (b) State x_2 with $N = 10^2$.

We assume the initial distribution of x_0 is $u_0(x) = \exp(-\frac{1}{2}|x|^4)$ and at each time step, the updated initial data are

$$u_i(x, \tau_i) = \exp\left\{(x_1[1 + 0.2\cos(x_2)], x_2[1 + 0.2\cos(x_1)])^\top dy(t)\right\} \cdot u_{i-1}(x, \tau_i).$$

The total simulation time is $T = 50$ s, and the update time step is $dt = 0.01$ s.

In Figs. 3 and 4, both of them are the one realization of the states x_1, x_2, respectively, in the 20 Monte Carlo simulations. They illustrate that our algorithm tracks the states better than the PF with 100 particles. The MSEs of our algorithm are about 0.64 and 0.52 on state x_1 and x_2, respectively, when the number of basis is $N = 5^2$. While the number increases to $N = 10^2$, the MSEs of our algorithm become relatively smaller, which are about 0.63 and 0.49 on state x_1 and x_2, respectively. The MSEs of x_1 by our algorithm with $N = 5^2$ and $N = 10^2$ also illustrate the proper number of the generalized Legendre polynomials should be chosen in the application of our algorithm. Whereas Table II is shown, in view that for the almost linear filtering systems, the EKF can be thought as the almost optimal filter, a benchmark. From the MSEs in Table II, we see that our algorithm in accuracy is nearly as well as that of EKF. As to the efficiency, the on-line CPU times of our algorithm is only 0.52 and 0.93s for $N = 5^2$ and 10^2, respectively.

TABLE II
PERFORMANCE OF DIFFERENT METHODS IN TIME-VARYING ALMOST LINEAR PROBLEM

Methods	CPU Time	MSE of state x_1	MSE of state x_2
Our algorithm with $N = 5^2$	0.52	0.6393	0.5250
Our algorithm with $N = 10^2$	0.93	0.6337	0.4983
EKF	0.24	0.6112	0.4632
PF	89.44	1.6214	1.0724

N means the total number of the elements in the set $V^2_{\sqrt{N}-1}$.

V. CONCLUSION

In this article, we explored the LSM to numerically solve the forward Kolmogorov equation (FKE) by adopting the generalized Legendre polynomial, which takes the vanishing boundary

condition into consideration in the spectral methods. Furthermore, the convergence rate of LSM to FKE is twice faster than the HSM in [29], under the suitable conditions. Equally importantly, our algorithm outperforms the EKF and PF in terms of estimation accuracy and efficiency in the numerical experiments of NLF problems. The experimental results illustrated the feasibility and efficiency of our proposed algorithm in NLF problems no matter in time-invariant or time-varying case. The complexity analysis of our approach and its application in much higher dimensional nonlinear systems are one of our future research topics.

Acknowledgment

The author Stephen S.-T. Yau would like to thank the National Center for Theoretical Sciences (NCTS) for providing excellent research environment while part of this articles was done.

References

[1] R. A. Adams, *Sobolev Spaces*. New York, NY, USA: Academic, 1978.
[2] J. D. Atkinson, "Eigenfunction expansions for randomly excited non-linear systems," *J. Sound Vib.*, vol. 30, no. 2, pp. 153–172, 1973.
[3] J. A. Anderson and S. L. Anderson, "A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts," *Monthly Weather Rev.*, vol. 127, no. 12, pp. 2741–2758, 1999.
[4] J. L. Anderson, "An ensemble adjustment Kalman filter for data assimilation," *Monthly Weather Rev.*, vol. 129, no. 12, pp. 2884–2903, 2001.
[5] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
[6] A. Bensoussan, R. Glowinski and A. Rascanu, "Approximation of the Zakai equation by the splitting up method," *SIAM J. Control Optim.*, vol. 28, pp. 1420–1431, 1990.
[7] A. Bensoussan, R. Glowinski, and A. Rascanu, "Approximation of some stochastic differential equations by the splitting up methods," *Appl. Math. Optim.*, vol. 25, pp. 81–106, 1992.
[8] A. Bain and D. Crisan, "Fundamentals of stochastic filtering," in Stochastic Model. Appl. Probability, vol. 60, New York, NY, USA: Springer-Verlag, 2009.
[9] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods in F Dynamics*. New York, NY, USA: Springer, 1988.
[10] T. Duncan, "Probability density for diffusion processes with applications to nonlinear filtering theory," Ph.D. dissertation, Stanford Electronics Labs, Stanford Univ., Standford, CA, USA, 1967.
[11] M. H. A. Davis and S. I. Marcus, "An introduction to nonlinear filtering," *in Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, M. Hazewinkel and J. S. Williams, Eds., Dordrecht, the Netherlands: Reidel, 53–75.
[12] S. Dolgov, B. N. Khoromskij, and I. V. Oseledets, "Fast solution of parabolic problems in the tensor train/quantized tensor train format with initial application to the Fokker-Planck equations," *SIAM J. Scientific Comput.*, vol. 33, no. 6, pp. A3016–A3038, 2012.
[13] G. Evensen, "The ensemble Kalman filter: Theoretical formulation and practical implementation," *Ocean Dyn.*, vol. 53, no. 4, pp. 343–367, 2003.
[14] Y. P. Chen, N. Y. Yi, and W. B. Liu, "A Legendre-Galerkin spectral method for optimal control problems governed by elliptic equations," *SIAM J. Numer. Anal.*, vol. 46, no. 5, pp. 2254–2275, 2008.
[15] D. Gottlieb and S. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*. , Philadelphia, PA, USA: SIAM, 1977.
[16] R. Ghanem and H. Sissaoui, "A posteriori error estimate by a spectral method of an elliptic optimal control problems," *J. Comput. Math. Optim.*, vol. 2, pp. 111–125, 2006.
[17] J. P. Johnson and R. A. Scott, "Extension of eigenfunction-expansion solutions of a fokker-planck equation-I. First order system," *Int. J. Non-Linear Mechanics*, vol. 14, pp. 315–324, 1979.
[18] J. P. Johnson and R. A. Scott, "Extension of eigenfunction-expansion solutions of a Fokker-Planck equation-II. Second order system," *Int. J. Non-Linear Mechanics*, vol. 15, pp. 41–56, 1980.
[19] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," in *Proc. IEEE*, 2004, vol. 92, no. 3, pp. 401–422, Mar. 2004.
[20] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.
[21] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *J. Basic Eng.*, vol. 83, pp. 95–108, 1961.
[22] H. Kushner, "Dynamical equations for optimal nonlinear filtering," *J. Differ. Equ.*, vol. 3, pp. 179–190, 1967.
[23] 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," *J. Sound Vib.*, vol. 237, no. 1, pp. 144–162, 2009.
[24] M. Kumar, S. Chakravorty, and J. L. Junkins, "On the curse of dimensionality in the Fokker-Planck equation," *in Proc. ASS Astrodyn. Specialist Conf.*, Pittsburgh, PA, USA, 2009, pp. 1781–1800..
[25] Z. G. Liang, S. S.-T. Yau, and S. T. Yau, "Finite-dimensional filters with nonlinear drift V: solution to Kolmogorov equation arising from linear filtering with non-Gaussian initial condition," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 4, pp. 1295–1308, Oct. 1997.
[26] S. Lakshmivarahan and D. J. Stensrud, "Ensemble Kalman filter application to meteorological data assimilation," *IEEE Control Syst. Mag.*, vol. 29, no. 3, pp. 34–46, Jul. 2009.
[27] S. Lototsky, R. Mikulevicius, and B. Rozovskii, "Nonlinear filtering revisited: a spectral approach," *SIAM J. Control Optim.*, vol. 82, pp. 35–45, 1960.
[28] X. Luo and S. S.-T. Yau, "Complete real time solution of the general nonlinear filtering problem without memory," *IEEE Trans. Autom. Control*, vol. 58, no. 10, pp. 2563–2578, Oct. 2013.
[29] X. Luo and S. S.-T. Yau, "Hermite spectral method to 1D forward Kolmogorov equation and its application to nonlinear filtering problems," *IEEE Trans. Autom. Control*, vol. 58, no. 10, pp. 2495–2507, Oct. 2013.
[30] X. Luo and F. Wang, "Generalized Jacobi spectral method in solving nonlinear filtering problems," *in Proc. IEEE Conf. Decis. Control*, Miami Beach, FL, USA, 2018, pp. 7206–7212.
[31] R. Mortensen, "Optimal control of continuous time stochastic systems," Ph.D. dissertation, Berkeley Electronics Research Lab., California Univ., Berkeley, CA, USA, 1996.
[32] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 34, no. 5, pp. 2295–2317, 2011.
[33] I. V. Oseledets and S. Dolgov, "Solution of linear systems and matrix inverson in the TT-format," *SIAM J. Sci. Comput.*, vol. 34, no. 5, pp. A2718–A2739, 2012.
[34] J. B. Qiu, K. K. Sun, T. Wang, and H. J. Gao, "Observer-based fuzzy adaptive event-triggered control for pure-feedback nonlinear systems with prescribed performance," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 11, pp. 2152–2162, Nov. 2019.
[35] J. Reif and R. Barakat, "Numerical solution of the Fokker-Planck equation via Chebyschev polynomial approximations with reference to first passage time probability density functions," *J. Comput. Phys.*, vol. 23, no. 4, pp. 425–445, 1977.
[36] H. Risken, H. D. Vollmer, and H. Denk, "Calculation of eigenvalues for the Karmers equation," *Phys. Lett. A*, vol. 78, no. 1, pp. 22–24, 1980.
[37] B. R. Jr., Spencer and L. A. Bergman, "On the numerical solution of the Fokker-Planck equation for nonlinear stochastic systems," *Nonlinear Dyn.*, vol. 4, pp. 357–372, 1993.
[38] J. Shen, "Efficient spectral-Galerkin method I. Direct solvers of second- and fourth-order equations using Legendre polynomials," *SIAM J. Sci. Comput.*, vol. 15, pp. 1489–1505, 1994.
[39] Y. Sun and M. Kumar, "A meshless p-PUFEM Fokker-Planck equation solver with automatic boundary condition enforcement," *in Proc. Amer. Control Conf.*, 2012, pp. 74–79.
[40] Y. Sun and M. Kumar, "Numerical solution of high dimensional stationary Fokker-Planck equation via tensor decomposition and Chebyshev spectral differential," *Comput. Math. Appl.*, vol. 67, no. 10, pp. 1960–1977, 2014.
[41] Y. Sun and M. Kumar, "A tensor decomposition method for high dimensional Fokker-Planck equation modeling polymeric liquids," in *Proc. 53rd AIAA Aerosp. Sci. Meeting*, Kissimmee, FL, USA, 2015, pp. 1–11.
[42] K. K. Sun, S. S. Mou, J. B. Qiu, T. Wang, and H. J. Gao, "Adaptive fuzzy control for nontriangular structural stochastic switched nonlinear systems with full state constraint," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 8, pp. 1587–1601, Aug. 2019.
[43] S. F. Wojtkiewicz, L. A. Bergman, and B. F. Jr. Spencer, "Numerical solution of some three state random vibration problems," *in Proc. 15th Biennial Conf. Mech. Vib. Noise*, Boston, MA, USA, 1995, pp. 939–947.

[44] S. F. Wojtkiewicz, L. A. Bergman, and B. F. Jr. Spencer, "High fidelity numerical solution of the Fokker-Planck equation," *in Proc. Int. Conf. Struct. Saf. Rel.*, Kyoto, Japan, 1998, pp. 933–940.
[45] S. F. Wojtkiewicz, E. A. Johnson, L. A. Bergman, M. Grigoriu, and B. F. Jr, Spencer, "Response of stochastic dynamical systems driven by additive Gaussian and Poisson white noise: Solution of a forward generalized Kolmogorov equation by a spectral finite difference method," *Comput. Methods Appl. Mech. Eng.*, vol. 168, pp. 73–89, 1999.
[46] W. V. Wedig and U. von Wagner, "Extended Laguerre polynomials for nonlinear stochastic systems," *in Computational Stochastic Mechanics*, P. D. Spanos, Ed., Rotterdam, Netherlands: A. A. Balkema, 1999, pp. 293–298.
[47] U. von Wagner and W. V. Wediger, "On the calculation of stationary solutions of multi-dimensional Fokker-Planck equations by orthogonal functions," *Nonlinear Dyn.*, vol. 21, pp. 289–306, 2000.
[48] S. T. Yau and S. S.-T. Yau, "Real time solution of nonlinear filtering problem without memory II," *SIAM J. Control Optim.*, vol. 47, no. 1, pp. 163–195, 2008.
[49] M. Zakai, "On the optimal filtering of diffusion processes," *Z. Wahrsch. Verw.* Gebiete, vol. 11, pp. 230–243, 1969.

Wenhui Dong was born in 1993 in Hubei, China. She received the Ph.D. degree in mathematics from Tsinghua University, Beijing, China, in 2020.

She is currently a Postdoctoral Fellow with the School of Mathematics and Information Science, Shaanxi Normal University, Xian, China. Her research interests include nonlinear filtering, estimation algebras, numerical analysis.

Xue Luo (Senior Member, IEEE) received the first Ph.D. degree in mathematics from East China Normal University (ECNU), Shanghai, China, in 2010 and the second Ph.D. degree in applied mathematics from the University of Illinois at Chicago (UIC), Chicago, IL, USA, in 2013.

During her study as a Ph.D. candidate with ECNU, she visited the Department of Mathematics, University of Connecticut, Storrs, CT, USA, during 2008–2009 and the Department of Mathematics, Statistics and Computer Science, UIC, during 2009–2010, as a Visiting Scholar, respectively. After her graduation from UIC, she joined Beihang University, Beijing, China. She is currently an Associate Professor with the School of Mathematical Sciences, BUAA. Her research interests include nonlinear filtering theory, numerical analysis of spectral methods, analysis of partial differential equations, sparse grid algorithm, and fluid mechanics.

Stephen S.-T. Yau (Fellow, IEEE) received the Ph.D. degree in mathematics from the State University of New York, Stony Brook, NY, US, in 1976.

He was a member of Institute of Advanced Study at Princeton during 1976–1977 and 1981–1982, and a Benjamin Pierce Assistant Professor with Harvard University, Cambridge, MA, USA, during 1977–1980. After that, he joined the Department of Mathematics, Statistics and Computer Science (MSCS), University of Illinois at Chicago (UIC), Chicago, IL, USA, and served for over 30 years. During 2005–2011, he became a Joint-Professor with the Department of Electrical and Computer Engineering and MSCS, UIC. After his retirement in 2012, he joined Tsinghua University, Beijing, China, where he is a Full-Time Professor with the Department of Mathematical Science. His research interests include nonlinear filtering, bioinformatics, complex algebraic geometry, CR geometry, and singularities theory.

Dr. Yau was the recipient of the Sloan Fellowship in 1980, Guggenheim Fellowship in 2000, IEEE Fellow Award in 2003, and AMS Fellow Award in 2013. In 2005, he was entitled the UIC Distinguished Professor. He is the Managing Editor and Founder of *Journal of Algebraic Geometry* from 1991, and the Editors-in-Chief and Founder of *Communications in Information and Systems* from 2000 till now. He was the General Chairman of IEEE International Conference on Control and Information, which was held in the Chinese University of Hong Kong in 1995.