Recent Progresses on Two Suboptimal Methods for Nonlinear Filtering Problems

by Xue Luo[*†] and Huimin Miao[‡]

Abstract. The nonlinear filtering (NLF) aims to yield a good estimation of the signal/state corrupted with noise, based on the noisy observations. In 2014's survey paper [31], the NLF methods are classified into two categories, the local and global approaches, by examining whether it approximates the posterior distribution of the states or only a finite number of the statistical quantities. Compared with the global approaches, the local ones are more computational friendly. In this survey, we shall discuss two recently developed suboptimal local methods for solving NLF problems, with emphasis on their reasonableness from a mathematical point of view.

Introduction

The study of nonlinear filtering (NLF) problems has a long history, tracing back to the work of Wiener [43] and Kolmogorov [26] in 1940s. The ultimate goal of filtering is to obtain "good" estimation of the states, a hidden stochastic process, based on the noisy observations. The states can be various quantities interested in different fields, e.g. the position and velocity of the moving targets in the tracking problem [28], the volatility process in finance [15], the locations of the hurrican in the atmospheric data assimilation problem [12], etc.

To be more specific, let us take the reentry problem, section 8.B, [21], as an example to illustrate the potential application of the filtering techniques. This model describes the dynamics of the vehicle from the space back to the earth. The reentry vehicle state of three-dimension is

$$x^T = [v, \gamma, \psi, h, \varphi, \theta], \tag{1}$$

where the first three components in x are the vehicle's velocity and angles, while the last three ones are the spherical coordinates of the vehicle. To be more specific, v is velocity (magnitude), γ is the flight path angle measured from the local horizontal, ψ is the heading angle measured from east toward north in the horizon plane, h is the altitude, φ the latitude, and θ the longitude. In this coordinate system, the motion of the vehicle is described by the differential equations

$$\begin{cases} \dot{v} = \dfrac{F_{xv}}{mv} - \dfrac{\mu}{r^2}\sin\gamma + \triangle\dot{v}_{rot}, \\ \dot{\gamma} = \dfrac{F_{zv}}{mv} + \dfrac{v\cos\gamma}{r} - \dfrac{\mu}{vr^2}\cos\gamma + \triangle\dot{\gamma}_{rot}, \\ \dot{\psi} = \dfrac{F_{yv}}{mv} - \dfrac{v}{r}\cos\gamma\cos\psi\tan\varphi + \triangle\dot{\psi}_{rot}, \\ \dot{h} = v\sin\gamma, \\ \dot{\varphi} = \dfrac{v}{r}\cos\gamma\sin\psi, \\ \dot{\theta} = \dfrac{v}{r\cos\varphi}\cos\gamma\cos\psi, \end{cases} \tag{2}$$

[*] School of Mathematical Sciences, Beihang University, Beijing, 102206, P. R. China.
E-mail: xluo@buaa.edu.cn
[†] Key Laboratory of Mathematics, Informatics and Behavioral Semantics (LMIB), Beihang University, Beijing, 100191, P. R. China.
[‡] School of Mathematical Sciences, Beihang University, Beijing, 102206, P. R. China.
E-mail: by1909011@buaa.edu.cn

where μ is the earth's gravitational constant, m is the vehicle mass, $r = R_E + h$, R_E is the radius of the earth, F_{xv}, F_{yv}, F_{zv} are the aerodynamic forces, and $\triangle \dot{v}_{rot}$, $\triangle \dot{\gamma}_{rot}$, $\triangle \dot{\psi}_{rot}$ are due to the earth's rotation. Intuitively, the first three equations in (2) describes the dynamics of the vehicle according to the Newton's law, while the last three equations give the change of the spherical coordinates during the vehicle's flight. The uncertainty in the states' evolution above comes from our relatively poor knowledge of the atmosphere (density, temperature, winds, etc.). Thus, the aerodynamic forces are not well known. The observations normally consist of ground-based radar tracking data (range, azimuth, and elevation) and telemetered onboard acceleration measurements. Taking the ground-based radar as an example, the range ρ is observed and given by

$$\rho = \left[(x-x_s)^2 + (y-y_s)^2 + (z-z_s)^2\right]^{1/2}, \tag{3}$$

where (x,y,z) is the vehicle's position and (x_s,y_s,z_s) is the position of the radar station. Thus, ρ is a nonlinear function of h, φ and θ. We refer the readers interested in the modelling of the reentry problem to [41]. Certainly, these observation data contain some information of the state we want, but inevitably polluted by instrumental noises. Therefore, the filtering techniques are needed to "filter" the state x in (1) out from the noisy observation data. In general, the expectation conditioned on the observation history is obtained by filtering. Sometimes, even the conditional covariance or the conditional distribution of the state are obtained depending on what filtering method one used.

In the literature, the filtering problems are classified as discrete, continuous and continuous-discrete ones [21, Chapter 5]. If both state and observation in the filtering problem are stochastic processes, i.e., continuous in time, like (2) and (3) in the reentry problem, then it is called continuous filtering problem. Similarly, if both state and observation are discrete in time, it is called discrete filtering problem. Analogously, the continuous-discrete filtering problems refer to those with continuous state and discrete observation. In mathematical modelling, the stochastic process is usually modelled by an Itô stochastic differential equation (SDE), while the discrete one is by a stochastic difference equation.

Let us take the continuous filtering problem as an example, it is modelled by the Itô SDE:

$$\begin{cases} dx_t = f(x_t,t)dt + g(x_t,t)dv_t, \\ dy_t = h(x_t,t)dt + dw_t, \end{cases} \tag{4}$$

where $x_t \in \mathbb{R}^d$ is the state process, $y_t \in \mathbb{R}^m$ is the observation process, f, g and h are drift, diffusion and observation function, respectively. The noises in the state and observation processes are modelled by two independent Brownian motions v_t and w_t. Notice that without the v_t and w_t terms, the equation of x_t in (4) becomes an ordinary differential equation (ODE), so that the trajectory of the state is deterministic, and the observation of the state is y_t. Without noises, it is not a filtering problem any more. Furthermore, the uncertainty of a Brownian motion is described by its covariance matrix. In the filtering literature, one usually assumes that $\mathbb{E}[dw_t dw_t^T] = Qdt$ and $\mathbb{E}[dv_t dv_t^T] = R(t)dt$, $R(t) > 0$, for all $t \in [0,T]$, respectively. The positive definite of $R(t)$ is a standard assumption, which asserts that the noise in the observation process won't vanish at any time which is called non-degenerate observation noise in the literature. The NLF problem with degenerate observation noise is another more difficult story, which has only been recently investigated under very special assumptions in [14]. In this survey, we shall make the standard non-degeneracy assumption. In the reentry problem above, the state to be estimated is x in (1). The nonlinear function of x on the right-hand side of (2) is denoted as f in (4). Meanwhile, the observation data $y_t = \rho$ is the range from the radar station in (3). In the linear case, i.e. f and h are both linear functions of x_t, and g is independent of x_t, the resulting ODE that the conditional expectation and the conditional covariance satisfy was derived by Kalman and Bucy [25]. This leads to an efficient algorithm for the continuous linear filtering problem. In fact, the discrete filtering problem appeared earlier than the continuous one. Already Gauss was interested in determining the orbital elements of a celestial body from observations. The influential Kalman filter (KF) [24] is the discrete version of Kalman-Bucy filter, which achieves the optimal estimation of the state in linear filtering problem in various sense, say minimum variance, maximum likelihood. Though KF was published one year earlier than the Kalman-Bucy filter, it was actually developed a little later.

In the nonlinear case, such efficient algorithms, as KF and Kalman-Bucy filter, has not been discovered so far due to its infinite-dimensional nature. Here, the dimension refers to the smallest number of statistical quantities needed to completely determine the conditional distribution of the state. For example, the dimension of normal distribution is two, since the distribution is completely determined by the expectation and the covariance. In fact, Hazewinkel et al. [19] took the cubic sensor problem as an example and pointed out that there exist no finite statistical quantities that can completely determine the state. At the International Congress of Mathematics in 1983, Brockett proposed to classify all finite-dimensional NLF problems, for which one may come up with an

efficient recursive algorithm similarly as the KF. It appears that only particularly structured classes of NLF problems bear computationally efficient recursive algorithms, e.g. Beneš filter [4] and Yau filter [8]. Starting from the Yau filter, a series of works conducted by Stephen S.-T. Yau and his co-workers characterize the structures of these classes from the viewpoint of Lie algebra, such as [44], [10], [47], [45], [48], [40], [22]. However, these structures are too special to yield an efficient recursive algorithm for general NLF problems.

Since it is known that the general NLF problems are essentially infinite dimensional, it is better to approximate the conditional distribution of the state

$$\mathbb{P}(A|\mathcal{Y}_t) = \int_{x\in A} p(x_t|\mathcal{Y}_t)dx, \tag{5}$$

with $A \in \mathbb{R}^d$ being a Borel set, and $\mathcal{Y}_t = \sigma\{y_s : \ s \leq t\}$, the filtration generated by the observations, instead of its finite-dimensional statistical quantities, e.g. the conditional expectation, the conditional covariance, etc. Already in the 1960s, Kushner [27] derived the Itô stochastic partial differential equation (SPDE), nowadays called Kushner equation, that the distribution of the states conditioned on the observation history satisfies. However, the Kushner equation is not computationally friendly, due to the nonlocal term involved. Duncan [11], Mortensen [38] and Zakai [52] independently came up an idea of viewing the conditional distribution under a new probability measure using the Girsanov transformation, just like the change of coordinates we always do in the Calculus. Under this new probability measure, the conditional distribution in Kushner equation become the so-called unnormalized conditional distribution, which satisfies a more favorable linear Itô SPDE, the DMZ equation. In [31], the algorithms that directly approximate the posterior density function are called global approaches, while those that only approximate some statistical quantities are local approaches. Once Kushner or DMZ equation has been derived, great efforts have been devoted to solve them numerically or analytically, for example the splitting-up method [5], the S^3 algorithm [29], and the direct method [46], [49], [9], [39], etc. We refer the interested readers to the survey [18] and references therein. But, these algorithms are not computationally efficient. Especially, the splitting-up algorithm proposed for the DMZ equation by Bensoussan et al. [5] behaves like the Trotter product formula from semigroup theory, which requires the boundedness of the drift and diffusion terms up to the second order to obtain the theoretical convergence analysis. Besides its theoretical limitations, there is no implementable scheme proposed. Until 2008 where Yau and Yau [50] published a pioneering work on the first feasible algorithm based on solving the robust piecewise DMZ equation. The original idea is that the heavy computation of solving the DMZ equation can be decomposed into two parts: the one is to solve Kolmogorov forward equation (KFE) off-line, and the other one is to synchronize the newly-coming observations online. Here, off-line refers to the computations made without any realization-dependent data. In the filtering, the observation data are realization-dependent and only available in each run of the simulation. After the decomposition, the heavy computational burden has been moved to the off-line part, which can be pre-computed and stored for later use. Since the on-line computation affects the computational efficiency. The real-time performance of the algorithm is foreseeable. It is inaccurate that the great contribution of the Yau-Yau method has been ignored and the credit was given to the splitting-up method in [34]. Later, the Yau-Yau algorithm has been validated by Luo et al. [34] for the time-varying system, and numerically verified the real-time performance when the state is one dimensional [35]. For the high-dimensional NLF problems, the numerical scheme solving the DMZ equation has been investigated in [36], [51] and [42].

Although the solutions to the Kushner or DMZ equations provide complete statistical information, they are highly computationally demanding, especially when the state is of high dimension. Efforts have been devoted in these global approaches to alleviate the "curse of dimensionality", for example by the sparse grid algorithm [36] or the proper decomposition method [42]. Compared to these global approaches, the local ones are much cheaper at the expense of obtaining only partial information of the states. To deal with the NLF problem, many variants of KF have been developed, including the extended Kalman filter (EKF) [16], the Gaussian sum filter (GSF) [1], the unscented Kalman filter (UKF) [23], the ensemble Kalman filter (EnKF) [13], the Gaussian quadrature Kalman filter (GKF) [20] and the cubature Kalman filter (CKF) [3, 2]. Unlike the KF in the linear filtering, all the variants of KF in the NLF problems are in general not optimal in any statistical sense, e.g. minimum variance or maximum likelihood. We thus refer to them as "suboptimal". In 2007, Germani et al. [17] used the Carleman approximation to form a bilinear system, for which the suboptimal estimation has been developed in [7]. To form a closed system, one has to ignore all higher moments. However, it may be inappropriate to do so for most NLF problems. Recently, the first author together with Jiao, Chiou and Yau developed two suboptimal methods for certain specific continuous nonlinear system [33, 37], based on the suboptimal filter for the bilinear system [7]. Under certain conditions, one can show that the higher order terms do converge to zero, as the order goes to infinity.

Of more practical use is the continuous-discrete NLF problems, where the measurements are sampled at discrete times. The first author together with Yau introduced a suboptimal linear estimation (see Definition 2) for the bilinear continuous-discrete system [32]. We first formulate the algorithm and then explain its "suboptimality".

In this survey, we shall summarize these suboptimal methods for the bilinear continuous/continuous-discrete system with emphasis on the reasonableness of these methods from a mathematical viewpoint.

Augmented States by Probabilists' Hermite Polynomials

In this section, we shall briefly introduce the improvement of the Carleman approach [17]. Its basic idea is very straightforward. Since the state satisfies a nonlinear system and is in general of infinite dimension, then one may believe that the more statistical quantities can be approximated recursively, the more accurate the estimation of the state is. Consequently, in the Carleman approximation [17], the SDE of the state's conditional moments was derived. To form a closed system, one has to truncate the infinite many moments at some order $\nu > 2$. However, this truncation is only justifiable if all the higher order conditional moments of the states in [17] are neglectable in some sense.

The first author together with Stephen S.-T. Yau proposed an alternative augmentation of the original state by its generalized Hermite polynomials (gHPs) $\left\{He_j^{\alpha,\beta}(x)\right\}_{j=0}^{\infty}$ (6) in [37]. With this alternation, one can show that $\mathbb{E}\left[He_j^{\alpha,\beta}(\xi)\right]$ tends to zero, as $j \to \infty$, if α, β are chosen appropriately and when the density function of the random variable ξ belongs to a particular class of functions. Thus, in contrast to the original Carleman approach, setting $He_j^{\alpha,\beta}(\xi) \equiv 0$, for all $j \geq \nu$, in the degree-ν approximation is better justified. Adjusting the two remaining parameters, the scaling factor α and the translating factor β, can lead to even better results.

For the clarity of the statements, we only write this method in the scalar case, i.e. $x_t \in \mathbb{R}$. The augmented state to the degree $\nu > 2$ is defined as

$$He_{1:\nu}^{\alpha,\beta}(x_t) = \left[He_1^{\alpha,\beta}(x_t)\ He_2^{\alpha,\beta}(x_t)\ \cdots\ He_\nu^{\alpha,\beta}(x_t)\right]^T.$$

For the high-dimensional state, the augmented states by the gHPs can be written using the Kronecker product:

$$\mathbf{He}^{\alpha,\beta}{}_{1:\nu} = \otimes_{i=1}^{d} He_{1:\nu,i}^{\alpha_i,\beta_i},$$

where $He_{1:\nu,i}^{\alpha_i,\beta_i} = \left[He_1^{\alpha_i,\beta_i}(x_i)\ He_2^{\alpha_i,\beta_i}(x_i)\ \cdots\ He_{\nu_i}^{\alpha_i,\beta_i}(x_i)\right]^T$, $i = 1,2,\cdots,d$. However, the numerical implementations of this suboptimal method may be inefficient compared to those in [17], due to the fact that the product of two Hermite polynomials (HPs) is no longer a HP, see (7).

Notations

1. The conditional expectation of $\circ$ is denoted by $\widehat{\circ} := \mathbb{E}[\circ \mid \mathcal{F}_t]$, where $\mathcal{F}_t := \sigma\{y_s : 0 \leq s \leq t\}$ is the observation history, modelled by a family of σ-algebra.
2. The Kronecker product $\otimes$ is used for the conciseness of the notations [6], and is defined for any two matrices $M_{r\times s}$ and $N_{p\times q}$:

$$M \otimes N := \begin{bmatrix} m_{11}N & \cdots & m_{1s}N \\ \cdots & \cdots & \cdots \\ m_{r1}N & \cdots & m_{rs}N \end{bmatrix}.$$

3. The gHP is defined as

$$He_n^{\alpha,\beta}(x) = He_n\left(\alpha(x-\beta)\right), \tag{6}$$

where $\{He_n(x)\}_{n=0}^{\infty}$ are the HP, $\alpha > 0$ is the scaling factor and $\beta \in \mathbb{R}$ is the translating factor. For any nonnegative integers n and m, we have

$$He_n^{\alpha,\beta}(x)He_m^{\alpha,\beta}(x) = \sum_{p\leq n\wedge m} A_{n,m,p} He_{n+m-2p}^{\alpha,\beta}(x), \tag{7}$$

where

$$A_{n,m,p} = \frac{n!m!}{p!(n-p)!(m-p)!}. \tag{8}$$

4. Given $\alpha > 0$ and $\beta \in \mathbb{R}$, any function $\circ \in C^{\infty}\left([0,T];L^2(\mathbb{R})\right)$ can be projected onto the gHPs:

$$\circ(x,t) = \sum_{k=0}^{\infty} \circ_k^{\alpha,\beta}(t) He_k^{\alpha,\beta}(x), \tag{9}$$

where $\circ_k^{\alpha,\beta}$ are smooth functions of t, which can be computed by

$$\circ_k^{\alpha,\beta}(t) = \frac{1}{\sqrt{2\pi}k!}\int_{\mathbb{R}} \circ(x,t) He_k{}^{\alpha,\beta}(x) e^{-\frac{\alpha^2(x-\beta)^2}{2}} dx. \tag{10}$$

Formulation

This suboptimal method is given in two steps: first, one gives the evolutionary SDE of the jth degree of the gHP's, $j = 0,1,\cdots$, and truncates the first ν degree of the gHP's equation to form a bilinear system; second, the similar procedure of KF has been applied to this bilinear system to yield a suboptimal estimation of the augmented states.

Step one: By Itô formula, one can easily derive the jth degree of the gHP's evolutionary SDE, $j = 0,1,\cdots$. Then writing the first νth degree of the gHPs

in the vector format, $He_{1:\nu}^{\alpha,\beta}(x_t)$, satisfies the following bilinear system:

$$\begin{cases} dHe_{1:\nu}^{\alpha,\beta}(x_t) = \left(\mathbf{F}_\nu He_{1:\nu}^{\alpha,\beta}(x_t) + \mathbf{F}_{\mathbf{0},\nu}\right) dt \\ \qquad\qquad + \left(\mathbf{G}_\nu He_{1:\nu}^{\alpha,\beta}(x_t) + \mathbf{G}_{\mathbf{0},\nu}\right) dv_t \\ dy_t = \left(H_\nu He_{1:\nu}^{\alpha,\beta}(x_t) + H_0\right) dt + dw_t \end{cases}, \tag{11}$$

where $\mathbf{F}_\nu$, $\mathbf{F}_{\mathbf{0},\nu}$, etc. are constant matrices of proper size. Thus, the NLF problem (4) has been approximated by a bilinear system, i.e. the drift, diffusion and observation functions are all linear functions of the state.

Step two: One obtain the evolutionary SDE of the suboptimal estimation of the augmented states $\{He_{1:\nu}^{\alpha,\beta}\}$ conditioned on the observation history [7].

Theorem 2.1. *For any given $\alpha > 0$ and $\beta \in \mathbb{R}$. $\widehat{He_{1:\nu}^{\alpha,\beta}}(x_t)$ satisfies the equation*

$$\begin{aligned} &d\widehat{He_{1:\nu}^{\alpha,\beta}}(x_t) \\ &= \left(\mathbf{F}_\nu \widehat{He_{1:\nu}^{\alpha,\beta}}(x_t) + \mathbf{F}_{\mathbf{0},\nu}\right) dt \\ &\quad + \left(\mathbf{G}_\nu m_\nu^{\alpha,\beta}(t) + \mathbf{G}_{0,\nu} + \mathbf{P}_\nu^{\alpha,\beta}(t) H_\nu^T\right) R^{-1} \\ &\quad \cdot \left[dy_t - \left(H_\nu \widehat{He_{1:\nu}^{\alpha,\beta}}(x_t) + H_0\right) dt\right], \end{aligned} \tag{12}$$

where $m_\nu^{\alpha,\beta} := \mathbb{E}\left(He_{1:\nu}^{\alpha,\beta}(x_t)\right)$ satisfy the following equations

$$\dot{m}_\nu^{\alpha,\beta}(t) = \mathbf{F}_\nu m_\nu^{\alpha,\beta}(t) + \mathbf{F}_{\mathbf{0},\nu} \tag{13}$$

with the initial values $m_\nu^{\alpha,\beta}(0) = \mathbb{E}\left(He_{1:\nu}^{\alpha,\beta}(x_0)\right)$, and $\mathbf{P}_\nu^{\alpha,\beta}(t)$ is the conditional error covariance matrix

$$\mathbf{P}_\nu^{\alpha,\beta}(t) = \mathbb{E}\left[\left(He_{1:\nu}^{\alpha,\beta}(x_t) - \widehat{He_{1:\nu}^{\alpha,\beta}}(x_t)\right)\right. \tag{14}$$

$$\left.\left(He_{1:\nu}^{\alpha,\beta}(x_t) - \widehat{He_{1:\nu}^{\alpha,\beta}}(x_t)\right)^T \middle| \mathcal{F}_t\right] \tag{15}$$

evolving according to the equation

$$\begin{aligned} \dot{\mathbf{P}}_\nu^{\alpha,\beta}(t) =& \mathbf{F}_\nu \mathbf{P}_\nu^{\alpha,\beta}(t) + \mathbf{P}_\nu^{\alpha,\beta}(t)\mathbf{F}_\nu^T + Q(t) \\ &- \left(\mathbf{G}_\nu m_\nu^{\alpha,\beta}(t) + \mathbf{G}_{0,\nu} + \mathbf{P}_\nu^{\alpha,\beta}(t) H_\nu^T\right) R^{-1} \\ &\cdot \left(\mathbf{G}_\nu m_\nu^{\alpha,\beta}(t) + \mathbf{G}_{0,\nu} + \mathbf{P}_\nu^{\alpha,\beta}(t) H_\nu^T\right)^T, \end{aligned} \tag{16}$$

with $\mathbf{P}_\nu^{\alpha,\beta}(0) = \Psi_\nu^{\alpha,\beta}(0)$.

Equation (12) and (16) give the evolution equations of the conditional expectation and the conditional covariance matrix, similarly as those in the KF for the linear filtering problem. The "gain function" (the second line on the right-hand side of (12)) plays the same role as the Kalman gain in the KF, to update the expectation by the innovation process, the difference between the real observation and the expected one (the last term inside the bracket in (12)). However, the difference between this suboptimal method and the KF is that this gain function depends not only the covariance matrix (14), but also the expectation without any information (13).

We remark that for $\nu = 1$, (12) and (16) coincide with the extended Kalman-Bucy filter (EKBF).

Mathematical Explanations

In this subsection, we shall explain from the mathematical point of view why this augmentation by the gHPs is possibly better than the conditional moments in [17]. This augmentation is motivated by a key observation [30]: $\mathbb{E}\left(He_j^{1,0}(\xi)\right) \to 0$, as $j \to \infty$, if α and β are appropriately chosen and the density function of the random variable ξ obeys Gaussian distribution. This observation can be easily extended to the scaled and translated gHPs, stated in the lemma below:

Lemma 1. *Suppose that the random variable $\xi \sim \mathcal{N}(a, b^2)$. Then for any $\mu \in \mathbb{R}$, we have*

$$\mathbb{E}\left[He_n^{1/b,a}(\xi + \mu)\right] = \left(\frac{\mu}{b}\right)^n.$$

In particular, for any $|\mu| < b$,

$$\lim_{n\to\infty} \mathbb{E}\left[He_n^{1/b,a}(\xi + \mu)\right] = 0. \tag{17}$$

This lemma shows that the higher order terms of the gHPs can be arbitrarily small, if the scaling and translating factors are chosen appropriately and $\xi \sim \mathcal{N}(a, b^2)$, and thus justifies the truncation in the bilinear system (11). But the assumption that $\xi \sim \mathcal{N}(a, b^2)$ is too restrictive to be of practical use for (11), as it implies that the conditional posterior distribution of the original state is Gaussian for all $t \in [0, T]$. This cannot be true unless (4) is actually a linear filtering. To weaken this assumption, the so-called exponential decay (ED) class of density functions is introduced.

Definition 1. *We say the density function $p(x) \in L^2(\mathbb{R})$ belongs to the exponential decay (ED) class with respect to (α, β), if for any $|\mu| < \frac{1}{\alpha}$, there exists some constant $C > 0$ and $\eta \in (0, 1 - \alpha|\mu|)$, such that*

$$|p_i| \le C\eta^i, \tag{18}$$

where p_i, $i = 0, 1, 2, \cdots$, are Fourier-Hermite coefficients of $p(x)$ defined in (10), i.e. $p_i = \frac{1}{\sqrt{2\pi i!}} \int_{\mathbb{R}} p(x,t) He_i^{\alpha,\beta}(x) e^{-\frac{\alpha^2(x-\beta)^2}{2}} dx$.

The ED class includes the Gaussian as a special case, with only the first two nonzero Fourier-Hermite

coefficients. The condition (18) is only on the Fourier-Hermite coefficients. But it actually reflects the regularity of the density by a similar argument as in the Riemann-Lebesgue lemma. The smoother the density is, the faster decay of the Fourier-Hermite coefficients to zero. The condition (18) implies the smoothness and the decay rate of the density, as $|x| \to \infty$. Consequently, combining Lemma 1 and Definition 1, we claim that by choosing the parameters α, β appropriately, the gHPs of high degree in (11) are small enough to be neglectable.

Theorem 2.2. *Let $\alpha > 0$ and $\beta \in \mathbb{R}$. If the random variable ξ has density function $p(x)$ belonging to the ED class with respect to (α, β), then for any $|\mu| < \frac{1}{\alpha}$ we have*

$$\lim_{n\to\infty} \left| \mathbb{E}\left[He_n^{\alpha,\beta}(\xi + \mu) \right] \right| = 0. \tag{19}$$

The assumption that the density belongs to the ED class can never be verified. Thus, there is always a probability that the suboptimal method cannot get a "good" estimation. This is a shortcoming from all the local approaches, no matter how delicate the design is. Moreover, the choice of the parameter α is crucial of the performance of the suboptimal method. We refer the interested readers to the detailed discussions on the implementation in [37, section 3.3 and 4].

Suboptimal Linear Estimation (SLE) for the Continuous-Discrete Bilinear System

As we explained, the NLF problem can be reduced to a bilinear system via Carleman approximation, see (11) in the previous section. In this section, we shall present a suboptimal linear estimation (SLE) for the continuous-discrete bilinear system, see the detailed discussions in [32].

The bilinear continuous-discrete system considered in the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is as follows:

$$\begin{cases} dX(t) = \mathbf{A}X(t)dt + \mathbf{N}dt + \sum_{j=1}^{b} (\mathbf{B}_j X(t) + \mathbf{F}_j) dW_j(t), \\ Y(t_k) = \mathbf{C}X(t_k) + \mathbf{D} + \sum_{j=1}^{b} \mathbf{G}_j V_j(t_k), \ k = 0, 1, \cdots, K, \end{cases} \tag{20}$$

for $t \in [0, T]$, where $0 = t_0 < t_1 < \ldots < t_K = T$, $\mathbf{A}$, $\mathbf{N}$, etc. are constant matrices of proper size. Similar as in (11), the drift, diffusion and observation functions are all linear in the state. Thus, (20) is a bilinear system, but with observations at discrete times. Moreover, $X(t) \in \mathbb{R}^n$ is the state, $Y(t_k) \in \mathbb{R}^m$ is the discrete observation data, $V_j(t_k) \sim \mathcal{N}(0, R_j(t_k)), R_j(t_k) \in \mathbb{R}$, $k = 0, 1, \ldots, K$, are independent one-dimensional white noises and $W_j(t), j = 1, \ldots, b$, are independent standard Brownian motions. Let $\mathscr{F}_{t_k}$ be the σ-field generated by the observations, i.e. $\mathscr{F}_{t_k} \triangleq \sigma\{Y(t_0), Y(t_1), \ldots, Y(t_k)\}$.

Throughout this section, we search for the linear estimation of $X(t_k)$. The linear estimation is defined in the sense that at each time instant t_k, the a priori estimate is a linear function of the posterior estimate, meanwhile the posterior estimate is the linear combination of the a priori estimate and the innovation as follows.

Definition 2. *We call $\hat{X}(t_k|t_k)$ the linear recursive estimate of $X(t_k)$ based on the observation $\{Y(t_0), Y(t_1), \ldots, Y(t_k)\}$, if*

1. *The a priori estimate, denoted as $\hat{X}(t|t_{k-1})$, $t \in [t_{k-1}, t_k]$, is linearly dependent of the posterior estimate $\hat{X}(t_{k-1}|t_{k-1})$ at time t_{k-1}, i.e.*

$$\hat{X}(t|t_{k-1}) = H_1(t)\hat{X}(t_{k-1}|t_{k-1}) + H_2(t), \tag{21}$$

   where H_1 and H_2 are matrices of proper dimensions;
2. *The posterior estimate $\hat{X}(t_k|t_k)$ lives in the linear space spanned by 1, the a priori linear estimate $\hat{X}(t_k|t_{k-1})$ and the innovation $Y(t_k) - \hat{Y}(t_k|t_{k-1})$, where $\hat{Y}(t_k|t_{k-1}) = \mathbf{C}\hat{X}(t_k|t_{k-1}) + \mathbf{D}$. More precisely,*

$$\begin{aligned} &\hat{X}(t_k|t_k) \\ &= H_3\hat{X}(t_k|t_{k-1}) + H_4(Y(t_k) - \hat{Y}(t_k|t_{k-1})) + H_5, \end{aligned} \tag{22}$$

   where H_3, H_4 and H_5 are constant matrices of proper dimensions.

In the section below, we shall give a two-step recursive algorithm for (20). It looks similar to the KF, which keeps track of the conditional mean and covariance matrix. The major difference comes from the nonlinearity of the system (20). The conditional posterior distribution of the state in the bilinear system is no longer always Gaussian. Thus, the first two central moments do not contain complete statistical information of the state. To keep the computational burden low, we still keep track of two statistical quantities in the algorithm, but with $Q(t|t_k)$ taking place of the covariance matrix $P(t|t_k)$. The substitute $Q(t|t_k)$ is given by a SDE (24) below. This replacement is meaningful, since $Q(t|t_k)$ is close to $P(t|t_k)$ in some sense by (S-2) below.

Formulation

The algorithm, we shall present below, consists of two steps: prediction and updating. The prediction is the propagation of the state according to the first equation in (20) between two observation times, while the updating is to synchronize the observation at each time instant t_k. This is a typical framework in the filtering literature.

Step 1: Prediction

In the interval $[t_{k-1},t_k)$, the a priori estimate $\hat{X}(t|t_{k-1})$ of $X(t)$ based on data $\{Y(t_0),Y(t_1),\ldots,Y(t_{k-1})\}$ satisfies

$$\dot{\hat{X}}(t|t_{k-1}) = \mathbf{A}\hat{X}(t|t_{k-1}) + \mathbf{N}, \tag{23}$$

$$\begin{aligned}\dot{Q}(t|t_{k-1}) &= \mathbf{A}Q(t|t_{k-1}) + Q(t|t_{k-1})\mathbf{A}^T \\ &+ \sum_{j=1}^{b}\Big[\mathbf{B}_j Q(t|t_{k-1})\mathbf{B}_j^T \\ &+ \left(\mathbf{B}_j\hat{X}(t|t_{k-1}) + \mathbf{F}_j\right)\left(\mathbf{B}_j\hat{X}(t|t_{k-1}) + \mathbf{F}_j\right)^T\Big],\end{aligned} \tag{24}$$

with the initial value $\hat{X}(t_{k-1}|t_{k-1})$ and $Q(t_{k-1}|t_{k-1})$ from previous updating, $\hat{X}(t_0|t_0) := \bar{X}_0$, and $Q(t_0|t_0) := \bar{P}_{t_0}$. Similar as in the KF, the evolution equations of the conditional expectation and the substitution of the conditional covariance are described by (23) and (24), respectively. It shows how the two statistical quantities of the state propagate between two observation times.

Step 2: Updating

The posterior estimate $\hat{X}(t_k|t_k)$ of $X(t_k)$ based on the observation history $\mathscr{F}_{t_k}$ satisfies

$$\hat{X}(t_k|t_k) = \hat{X}(t_k|t_{k-1}) + K_k\left[Y(t_k) - \hat{Y}(t_k|t_{k-1})\right], \tag{25}$$

with $\hat{Y}(t_k|t_{k-1}) = \mathbf{C}\hat{X}(t_k|t_{k-1}) + \mathbf{D}$, and the gain function K_k is given by

$$K_k = Q(t_k|t_{k-1})\mathbf{C}^T\left[\mathbf{C}Q(t_k|t_{k-1})\mathbf{C}^T + \sum_{j=1}^{b}\mathbf{G}_j R_j(t_k)(\mathbf{G}_j)^T\right]^{-1}. \tag{26}$$

Meanwhile, the matrix $Q(t_k|t_k)$ is updated by

$$Q(t_k|t_k) = (I_n - K_k\mathbf{C})\,Q(t_k|t_{k-1}), \tag{27}$$

where I_n is the identity matrix of dimension $n\times n$.

We remark that compared to the KF, the matrix $Q(t|t_{k-1})$ plays the role of the conditional variance

$$\begin{aligned}&P(t|t_{k-1}) \\ :=&\mathbb{E}\left[(X(t)-\hat{X}(t|t_{k-1}))(X(t)-\hat{X}(t|t_{k-1}))^T\,\middle|\,\mathscr{F}_{t_{k-1}}\right]\end{aligned}$$

in SLE. However, $Q(t|t_{k-1}) \neq P(t|t_{k-1})$ in general, $t\in(t_{k-1},t_k)$, even if they have the same initial value at $t = t_{k-1}$.

Mathematical Properties of the SLE

In this subsection, we shall detail the nice mathematical properties of the algorithm advertized in the previous section. The only assumption we made is that

(As) $\mathbf{A}$ and $\mathbf{A}_{ex}$ are Hurwitz, where $\mathbf{A}_{ex} := \sum_{l=1}^{b}(\mathbf{B}_l\otimes\mathbf{B}_l) + I_n\otimes\mathbf{A} + \mathbf{A}\otimes I_n$.

This assumption essentially guarantees the stability of the bilinear system. Under this assumption, the algorithm presented before gives indeed a suboptimal linear estimate (SLE) in the following sense:

(S-1) If $\hat{X}(t_0|t_0)$ is unbiased, so is $\hat{X}(t_k|t_k)$, i.e.,

$$\mathbb{E}\left(\hat{X}(t_k|t_k) - X(t_k)\right) = 0, \tag{28}$$

for $k = 1,\cdots,K$.

(S-2) The a priori estimate $\hat{X}(t|t_{k-1})$, for $t\in[t_{k-1},t_k]$ and $Q(t|t_{k-1})$ converge to the conditional expectation $\mathbb{E}\left(X(t)|\mathcal{F}_{t_{k-1}}\right)$ and the conditional variance $P(t|t_{k-1})$ component-wise and exponentially fast with respect to t.

(S-3) The posterior estimate $\hat{X}(t_k|t_k)$ minimizes the conditional variance error $\mathrm{tr}P(t_k|t_k)$ in the linear space spanned by $\{1,\hat{X}(t_k|t_{k-1}),Y(t_k)-\hat{Y}(t_k|t_{k-1})\}$, where $\mathrm{tr}(\circ)$ denotes the trace of $\circ$.

The proofs of these three properties (S-1)–(S-3) are involved. We refer the interested readers to [32]. We remark that

1. As for (S-1), the posterior estimate $\hat{X}(t_k|t_k)$ is a $\mathcal{F}_{t_k}$-measurable random variable, as the conditional expectation $\mathbb{E}(X_{t_k}|\mathcal{F}_{t_k})$. But they are not identical in general.
2. If the posterior estimate at t_{k-1} is the conditional expectation at t_{k-1}, then by (S-2) the estimate coincides with the conditional expectation for all $t\in[t_{k-1},t_k]$.
3. (S-3) tells us that $\mathrm{tr}P(t_k|t_k)$ is minimized by properly chosen K_k (26). And (S-2) asserts that $Q(t|t_{k-1})$ approaches to $P(t|t_{k-1})$ exponentially fast. Consequently, $Q(t_k|t_k)$ is "almost minimized" by K_k in (26), since $Q(t_k|t_{k-1})$ is close to $P(t_k|t_{k-1})$.

Acknowledgements

This work is financially supported by National Key R&D Program of China (grant no. 2022YFA100503), National Natural Science Foundation of China (grant no. 12271019, 11871003, 11961141005) and the Fundamental Research Funds for the Central Universities (grant no. YWF-22-L-640).

References

[1] D. Alspach and H. Sorenson. Nonlinear Bayesian estimation using Gaussian sum approximation. *IEEE Trans. Automat. Control*, 17(4):439–448, 1972. MR0321581

[2] I. Arasaratnam and I. Haykin, S.and Hurd. Cubature Kalman filtering for continuous-discrete systems: theory and simulations. *IEEE Trans. Signal Process.*, 58(10):4977–4993, 2010. MR2722659

[3] I. Arasaratnam and S. Haykin. Cubature Kalman filters. *IEEE Trans. Automat. Control*, 54(6):1254–1269, 2009. MR2532614

[4] V. E. Beneš. Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics*, 5(1-2):65–92, 1981. MR0643062
[5] A. Bensoussan, R. Glowinski, and A. Rascanu. Approximation of the Zakai equation by the splitting-up method. *SIAM J. Control Optim.*, 28:1420–1431, 1990. MR1075210
[6] F. Carravetta, A. Germani, and M. Raimondi. Polynomial filtering for linear discrete time non-Gaussian systems. *SIAM J. Control Optim.*, 34(5):1666–1690, 1996. MR1404851
[7] F. Carravetta, A. Germani, and M. Shuakayev. A new suboptimal approach to the filtering problem for bilinear stochastic differential systems. *SIAM J. Control Optim.*, 38(4):1171–1203, 2000. MR1760065
[8] J. Chen. On ubiquity of Yau filters. In *Proc. Amer. Control Conf.*, pages 252–254, Baltimore, MD, 1994.
[9] X. Chen, X. Luo, and S. S.-T. Yau. Direct method for time-varying nonlinear filtering problems. *IEEE Trans. Aerospace Electron. Systems*, 53(2):630–639, 2017.
[10] W. Chiou and S. S.-T. Yau. Finite-dimensional filters with nonlinear drift. II. Brockett's problem on classification of finite-dimensional estimation algebras. *SIAM J. Control Optim*, 32(1):297–310, 1994. MR1255972
[11] T. Duncan. *Probability densities for diffusion processes with applications to nonlinear filtering theory*. PhD thesis, Stanford University, United States, 1967.
[12] G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(5):10143–10162, 1994.
[13] G. Evensen. The ensemble Kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, 2003.
[14] D. J. Fotsa-Mbogne and E. Pardoux. Nonlinear filtering with degenerate noise. *Electron. Commun. Probab.*, 22(44):1–14, 2017. MR3693770
[15] R. Frey and W. J. Runggaldier. A nonlinear filtering approach to volatility estimation with a view towards high frequency data. *International Journal of Theoretical & Applied Finance*, 4(02):199–210, 2001. MR1831267
[16] A. Gelb. *Applied Optimal Estimation*. MIT Press, Cambridge, MA, 1974.
[17] A. Germani, C. Manes, and P. Palumbo. Filtering of stochastic nonlinear differential systems via a Carleman approximation approach. *IEEE Transactions on Automatic Control*, 52(11):2166–2172, 2007. MR2362165
[18] I. Gyongy. Approximations of stochastic partial differential equations. In *Lecture Notes in Pure and Applied Mathematics*, volume 277, pages 287–307, New York, 2002. Dekker. MR1919514
[19] M. Hazewinkel, S. I. Marcus, and H. J. Sussmann. Nonexistence of finite-dimensional filters for conditional statistics of the cubic sensor problem. *Syst. Contr. Lett.*, 3:331–340, 1983. MR0729114
[20] K. Ito and K. Xiong. Gaussian filters for nonlinear filtering problems. *IEEE Transactions on Automatic Control*, 45(5):910–927, 2000. MR1774134
[21] A. Jazwinski. *Stochastic processes and filtering theory*. Academic Press, New York, 1970.
[22] X. Jiao and S. S.-T. Yau. New classes of finite dimensional filters with nonmaximal rank estimation algebra on state dimension n and linear rank n-2. *SIAM J. Control Optim.*, 58(6):3413–3427, 2020. MR4175395
[23] S. Julier and J. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422, 2004.
[24] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME Ser. D. J. Basic Engrg.*, 82(1):35–45, 1960. MR3931993
[25] R. E. Kalman and R. Bucy. New results in linear filtering and prediction theory. *Trans. ASME Ser. D. J. Basic Engrg.*, 83(1):95–108, 1961. MR0234760
[26] A. Kolmogorov. Interpolation and extrapolation of stationary random sequences. *Izv. Akad. Nauk USSR, ser. Math*, 5(5):3–14, 1941.
[27] H. Kushner. On the differential equations satisfied by conditional probability densities of Markov processes, with applications. *Journal of SIAM Control Series A*, 2(1):106–119, 1964. MR0180407
[28] H. J. Kushner and A. S. Budhiraja. A nonlinear filtering algorithm based on an approximation of the conditional distribution. *IEEE Trans. Automat. Control*, 45(3):580–585, 2000. MR1762882
[29] S. Lototsky, R. Mikulevicius, and B. Rozovskii. Nonlinear filtering revisited: A spectral approach. *SIAM J. Control Optim.*, 35(2):435–461, 1997. MR1436632
[30] W. Luo. *Wiener chaos expansion and numerical solutions of stochastic partial differential equations*. PhD thesis, California Institute of Technology, United States, 2006. MR3078549
[31] X. Luo. On recent advance of nonlinear filtering theory: emphasis on global approaches. *Pure Appl. Math. Quart.*, 10:685–721, 2014. MR3324765
[32] X. Luo, X. Chen, and S. S.-T. Yau. Suboptimal linear estimation for continuous-discrete bilinear systems. *Systems & Control Letters*, 119:92–100, 2018. MR3842605
[33] X. Luo, Y. Jiao, W.-L. Chiou, and S. S.-T. Yau. A novel suboptimal method for solving polynomial filtering problems. *Automatica*, 62:26–31, 2015. MR3423968
[34] X. Luo and S. S.-T. Yau. Complete real time solution of the general nonlinear filtering problem without memory. *IEEE Trans. Automat. Control*, 58(10):2563–2578, 2013. MR3106062
[35] X. Luo and S. S.-T. Yau. Hermite spectral method to 1d forward Kolmogorov equation and its application to nonlinear filtering problems. *IEEE Trans. Automat. Control*, 58(10):2495–2507, 2013. MR3106057
[36] X. Luo and S. S.-T. Yau. Hermite spectral method with hyperbolic cross approximations to high-dimensional parabolic PDEs. *SIAM J. Numer. Anal.*, 51(6):3186–3212, 2013. MR3131484
[37] X. Luo and S. S.-T. Yau. The suboptimal method via probabilists' Hermite polynomials to solve nonlinear filtering problems. *Automatica*, 94:9–17, 2018. MR3810950
[38] R. Mortensen. *Optimal control of continuous time stochastic systems*. PhD thesis, University of California at Berkeley, United States, 1966.
[39] J. Shi, Z. Yang, and S. S.-T. Yau. Direct method for Yau filter system with nonlinear observations. *Internat. J. Control*, 91(3):678–687, 2018. MR3760421
[40] J. Shi and S. S.-T. Yau. Finite dimensional estimation algebras with state dimension 3 and rank 2, I: Linear structure of Wong matrix. *SIAM J. Control Optim.*, 55(6):4227–4246, 2017. MR3738843
[41] W. E. Wagner. Re-entry filtering, prediction and smoothing. *J. Spacecraft Rockets*, 3:1321–1327, 1966.
[42] Z. Wang, X. Luo, S. S.-T. Yau, and Z. Zhang. Proper orthogonal decomposition method to nonlinear filtering problems in medium-high dimension. *IEEE Trans. Automat. Control*, 65(4):1613–1624, 2020. MR4085532
[43] N. Wiener. *Extrapolation, Interpolation and Smoothing of Time Series, with Engineering Applications*. Wiley, New York, 1949. MR0031213
[44] S. S.-T. Yau. Finite-dimensional filters with nonlinear drift I: A class of filters including both Kalman-Bucy filters and Benes filters. *J. Math. Systems Estim. Control*,

4:181–203, 1994. MR1298555
[45] S. S.-T. Yau. Complete classification of finite-dimensional estimation algebras of maximal rank. *Internat. J. Control*, 76:657–677, 2003. MR1979888
[46] S. S.-T. Yau and G.-Q. Hu. Direct method without Riccati equation for Kalman-Bucy filtering system with arbitrary initial conditions. In *Proceedings of the 13th World Congress IFAC*, pages 469–474, San Francisco, 1996.
[47] S. S.-T. Yau and G.-Q. Hu. Finite-dimensional filters with nonlinear drift x: explicit solution of DMZ equation. *IEEE Trans. Automat. Control*, 46(1):142–148, 2001. MR1809477
[48] S. S.-T. Yau and G.-Q. Hu. Classification of finite-dimensional estimation algebras of maximal rank with arbitrary state-space dimension and mitter conjecture. *Internat. J. Control*, 78(10):689–705, 2005. MR2152444
[49] S. S.-T. Yau and Y. Lai. Explicit solution of DMZ equation in nonlinear filtering via solution of ode. *IEEE Trans. Automat. Control*, 48:505–508, 2003. MR1962262
[50] S.-T. Yau and S. S.-T. Yau. Real time solution of the nonlinear filtering problem without memory II. *SIAM J. Control Optim.*, 47(1):163–195, 2008. MR2373467
[51] M.-H. Yueh, W.-W. Lin, and S.-T. Yau. An efficient numerical method for solving high-dimensional nonlinear filtering problems. *Commun. Inf. Syst.*, 14:243–262, 2014. MR3340942
[52] M. Zakai. On the optimal filtering of diffusion processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 11(3):230–243, 1969. MR0242552