Complete Solution to the Most General Nonlinear Filtering Problems and Its Implementation

Stephen S.-T. Yau

Department of Mathematical Sciences,
Tsinghua University, Beijing, China.
&
Yanqi Lake Beijing Institute of Mathematical Sciences
and Applications, Beijing, China.

University of Notre Dame, October 27, 2022

Abstract

The famous filtering problem of estimating the state of a stochastic dynamical system from noisy observations is of central importance in engineering. This problem is reduced to solve the Duncan-Mortensen-Zakai (DMZ) equation which is satisfied by the unnormalized conditional density of the state given the observation history. We first introduce a very general class of finite dimensional filters which include Kalman filter and Benes filter as special cases. For general nonlinear filtering problems, solving the DMZ equation can be reduced to solving a Kolmogorov forward equation off-line, and this is the so-called Yau-Yau algorithm. At last, we shall introduce the filter based on the recurrent neural networks (RNNs).

Table Contents

1 Filtering problems
- Goal of filterings
- History
- Model and Duncan-Mortensen-Zakai (DMZ) equation

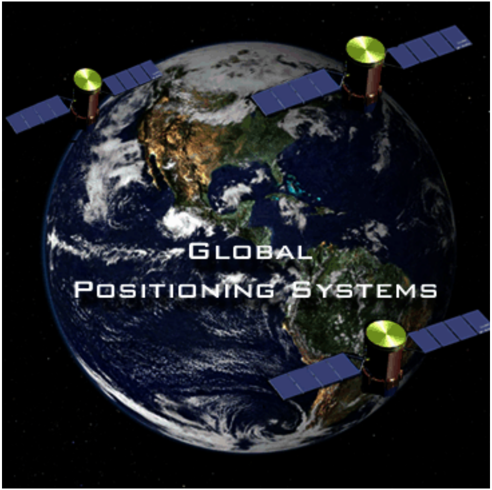
2 Finite dimensional filter
- Estimation algebra
- Direct method

3 Yau-Yau algorithm
- Validation of our algorithm
- Aim and assumptions
- Our alogrithm
- Experiments

4 Neural network filter

Goal of filterings

Goal: to form the "best estimate" for the true value of some system, given only some potentially noisy observations of that system.



Kalman filter and its applications

R. E. Kalman, 1960: Kalman filter – Optimal linear filter

Application:

- in the navigation of Apollo 13 – by providing the estimates of its trajectory to guide it to the Moon and back;
- in the navigation systems of U.S. Navy nuclear ballistic missile submarine;
- in the guidance and navigation systems of cruise missiles, such as the U.S. Navy's Tomahawk missile and the U.S. Air Force's Air Launched Cruise Missile;
- in the guidance and navigation systems of the NASA Space Shuttle and the International Space Station.

Award: Because of Kalman filter, R. E. Kalman is awarded Charles Stark Draper Prize – one of three prizes that constitute the "Nobel Prizes of Engineering".

Drawbacks of Kalman filter and its derivatives

R. E. Kalman and R. S. Bucy, 1961:
Kalman-Bucy filter – continuous time version of the Kalman filter

"They try all sorts of fixes, but basically the problem is such that the linear theory does not apply"

– R. S. Bucy, *SIAM News* **26**, Aug 1993

Failures of Kalman filter may due to

- Nonlinearity: the outputs are **not** a linear function of the inputs;
- Non-Gaussian of the initial states.

Even its **derivatives**, such as Extended Kalman filter, Unscented Kalman filter, Ensemble Kalman filter, etc can **NOT** avoid these two dead spots completely.

Nonlinear filterings (NLF)

Office of Naval Research (around 1995)

Given the noisy observation of the real states, can we give the "accurate" estimates of the states instantaneously, provided as much computational resources as one needs?

It has been an OPEN question for more than 50 years. It is finally SOLVED theoretically in this talk.

Attempts

Attempts without much success:

- V. E. Beneš, 1981: derives an exact filter for a special class of nonlinear problems, so-called Beneš filter;
  – Does **not** include all linear problems.
- Around 1980, S. Mitter and R. Brockett proposed to use Lie algebra method to solve NLF. Finite dimensional Lie algebra will give finite dimensional filter. In a series of papers, Yau with his various collaborators classify all finite dimensional nonlinear filters of maximal rank.

Popular NLFs

Widely used NLFs nowadays:

Existing filters	*Shortcomings*
Assumed-density filter (extended Kalman filter)	Nonlinearity
	Gaussian assumption of initial state
Sequential Monte Carlo methods (particle filter)	Can't be implemented in real time

Signal based model

We consider the following signal based 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: states, n-vector;
- f: drift term, n-vector;
- G: diffusion term, $n \times r$ matrix;
- y_t: observation path, m-vector;
- h: observation term, m-vector;
- v_t: r-vector Brownian motion with $E[dv_t dv_t^T] = Q(t)dt$;
- w_t: m-vector Brownian motion with $E[dw_t dw_t^T] = S(t)dt$ and $S(t) > 0$.

Assume $y_0 = 0$ and x_0, $\{v_t, t \geq 0\}$, $\{w_t, t \geq 0\}$ are independent.

Duncan-Mortensen-Zakai (DMZ) equation

1960s, Duncan, Mortensen and Zakai:

$\sigma(x,t)$: unnormalized density function of x_t conditioned on the observation history $Y_t = \{y_s : 0 \leq s \leq t\}$.

It satisfies the DMZ equation:

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

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

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

"Pathwise-robust" DMZ equation

Construct robust state estimators from any observed sample paths: For each "given" observation y_t, let (Rozovsky, 1972)

$$\sigma(x,t) = \exp\left[h^T(x,t)S^{-1}(t)y_t\right]u(x,t),$$

it yields the "pathwise-robust" DMZ equation:

$$\begin{cases} \dfrac{\partial u}{\partial t}(x,t) + \dfrac{\partial}{\partial t}(h^T S^{-1})^T y_t u(x,t) \\ \qquad\qquad = \exp\left(-h^T S^{-1} y_t\right)\left[L - \dfrac{1}{2}h^T S^{-1} h\right]\left[\exp\left(h^T S^{-1} y_t\right)u(x,t)\right] \\ u(x,0) = \sigma_0(x). \end{cases} \tag{4}$$

Approximation

Let u_i be the solution of the "pathwise-robust" DMZ equation with y_t freezed at $t = \tau_{i-1}$, for $\tau_{i-1} \leq t \leq \tau_i$, $i = 1, 2, \cdots, k$, with initial data

$$u_1(x, 0) = \sigma_0(x), \quad u_i(x, \tau_{i-1}) = u_{i-1}(x, \tau_{i-1}), \quad \text{for } i = 2, 3, \cdots, k.$$

Expect:

$$\hat{u}(x, t) = \sum_{i=1}^{k} \chi_{[\tau_{i-1}, \tau_i]}(t) u_i(x, t) \to u(x, t)$$

in some sense, as $|\mathcal{P}_k| = \sup_{1 \leq i \leq k}(\tau_i - \tau_{i-1}) \to 0$.

Introduction to estimation algebra

1970s: Brockett and Clark, Brockett, and Mitter proposed estimation algebras method

1983 (International Congress of Mathematics): Brockett proposed the problem of classifying finite dimensional estimation algebras (FDEA).

Advantages:

- It takes into account of geometrical aspects of the situation.
- As long as the estimation algebra is finite dimensional, the finite dimensional recursive filter can be constructed explicitly.
- Lie algebraic methods are highly useful for classifying equivalence of finite dimensional filters.
- The number of sufficient statistics in the Lie algebra method linearly depends on state space dimension.

Basic concept

If noises in state equation and observation equation are independent standard Brownian motions, i.e., $Q(t) = S(t) = I$, then we define

$$L_0 := \frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2} - \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i} - \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i} - \frac{1}{2}\sum_{i=1}^{m} h_i^2. \quad (5)$$

For $i = 1, \cdots, m$, L_i is defined the zero degree differential operator of multiplication by h_i.

Definition 1

The estimation algebra E of a filtering system (1) is defined to be the Lie algebra generated by $\{L_0, L_1, \cdots, L_m\}$, i.e., $E = \langle L_0, h_1, \cdots, h_m \rangle_{L.A.}$. Furthermore, if $f = \nabla\phi$ for some $\phi \in C^\infty(\mathbb{R}^n)$, E is called exact.

Basic concept

Definition 2

Let $L(E) \subset E$ be the vector space consisting of all the homogeneous degree 1 polynomials in E. Then the linear rank of estimation algebra E is defined by $r := \dim L(E)$. Especially, if $r = n$, we call E has maximal rank.

Based on the structure of linear rank, classifications of estimation algebra have always been a research hotspot. Especially, from 1990 to 1997, in the series work of Yau and coworkers[1][2][3], complete classification of maximal rank estimation algebra has been finished[4],.

[1] Chen and Yau, *Math. Control Signals Systems*, 1996
[2] Chiou and Yau, *SIAM J. Control Optim.*, 1994
[3] Yau, *J. Math. Systems Estim. Control*, 1994
[4] Yau, *Internat. J. Control*, 2003

Basic concept

In 1994, Yau[5] proposed a very useful theorem in underdetermined PDE appearing in estimation algebra.

Theorem 3 (Yau)

Let $F(x_1, \cdots, x_n)$ be a C^∞ function on R^n. Suppose that there exists a path $c : R \to R^n$ and $\delta > 0$ such that $\lim_{t\to\infty} \|c(t)\| = \infty$ and $\lim_{t\to\infty} \sup_{B_\delta(c(t))} F = -\infty$, where $B_\delta(c(t)) = \{x \in R^n : \|x - c(t)\| < \delta\}$. Then there are no C^∞ functions $f_1, f_2, \cdots, f_n$ on R^n satisfying the equation

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = F. \tag{6}$$

[5] Yau, *J. Math. Systems Estim. Control*, 1994

Structures of finite-dimensional exact estimation algebras

We shall survey some estimation algebra related results in 1990s. Let $\eta := \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{m} h_i^2$ and $D_i := \frac{\partial}{\partial x_i} - f_i,\ 1 \le i \le n$. Following result describes finite dimensionality of E by an algebraic condition[6].

Theorem 4 (Tam, Wong, & Yau)

Suppose E is an exact estimation algebra. Then E is finite-dimensional if and only if $\nabla h_i^T J_\eta^j$ is a constant for $1 \le i \le m$ and all $j = 0, 1, \cdots$ where J_η is the Hessian matrix of η.

Theorem 5 (Dong, Tam, Wong & Yau)

Suppose E is a finite-dimensional exact estimation algebras of maximal rank. Then it is a real vector space of dimension $2n+2$ with basis given by $1, x_1, x_2, \cdots, x_n, D_1, \cdots, D_n$ and L_0.

[6] Tam, Wong and Yau, *SIAM J. Control Optim.*, 1990

Estimation algebras of maximal rank with Ω-matrix in constant coefficients

For estimation algebra E whose Ω-matrix has constant entries, in 1994, Yau[7] proposed a sufficient condition for E to be finite dimensional:

$$\begin{cases} \deg(\eta) \leq 2 \\ \deg(h_i) \leq 1, i = 1, \cdots, m \end{cases} \implies \dim E < \infty \tag{7}$$

and conversely, if E is finite dimensional, then $h_1, \cdots, h_m$ are affine in x, i.e., the observation matrix $H = [\nabla h_1, \cdots, \nabla h_m]$ is a constant matrix. Furthermore, if the observation matrix has rank n, then η is a polynomial of degree at most 2 and E is of dimension $2n + 2$ with a basis given by $1, x_1, \cdots, x_n, D_1, \cdots, D_n, L_0$.

[7] Yau, *J. Math. Systems Estim. Control*, 1994

Wei-Norman approach

Robust DMZ equation is a time-varying PDE. Generally, we consider following PDE:

$$\frac{\partial u}{\partial t} = a_1(t)A_1 u + \cdots + a_m(t)A_m u, \tag{8}$$

where the A_i's are linear partial differential operators in $x_1, \cdots, x_n$, and the a_i's are given functions of time t. We shall assume that the Lie algebra generated by the operators A_i's is finite dimensional. Without loss of generality, we can assume A_i's consist a basis of Lie algebra.

The central idea of Wei-Norman theory is to find a solution of the form

$$u(t,x) = e^{g_1(t)A_1} \dots e^{g_m(t)A_m} \psi(x), \tag{9}$$

where the g_i's satisfy a system of ODEs and can be determined uniquely.

Classification of maximal rank estimation algebra

The programme of classifying FDEA of maximal rank was begun in 1990 by Yau. There are four crucial steps here. First we define following Wong matrix $\Omega = (\omega_{ij})$. where

$$\omega_{ij} = \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}, \quad \forall 1 \le i, j \le n.$$

Obviously, $\omega_{ij} = -\omega_{ji}$, i.e., Ω is an anti-symmetric matrix.
Step 1. In 1990, Yau first observed that Wong's Ω matrix plays an important role. As the first crucial step, he classifies all finite dimensional estimation algebras of maximal rank if Wong's matrix has entries in constant coefficients[8]. Later, Chiou and Yau[9] study maximal rank estimation algebra when state dimension $n \le 2$.

[8]Yau, *J. Math. Systems Estim. Control*, 1994
[9]Chiou and Yau, *SIAM J. Control Optim.*, 1994

Classification of maximal rank estimation algebra

Step 2. The second crucial step was due to Chen and Yau[10] in 1996. They developed quadratic structure theory. In particular, they introduced the notion of quadratic rank k. In this way, the Wong's Ω-matrix is divided into three parts:

(1) $(\omega_{ij}), 1 \leq i, j \leq k$;

(2) $(\omega_{ij}), k+1 \leq i, j \leq n$

(3) $(\omega_{ij}), 1 \leq i \leq k, k+1 \leq j \leq n$, or $1 \leq j \leq k, k+1 \leq i \leq n$.

Chen and Yau (1997) proved among many other things that part (1) is a matrix with constant coefficients.

[10] Chen and Yau, *SIAM J. Control Optim.*, 1997

Classification of maximal rank estimation algebra

Step 3. In 1997, Wu, Yau and Hu[11] first proved the weak Hessian matrix non-decomposition theorem for general n. Thus part (2) is also a constant matrix for arbitrary n. In 1999, Yau, Wu and Wong [12]dropped the restriction of cyclical condition and extended to strong Hessian nondecomposition theorem successfully.

Theorem 6 (Strong Hessian nondecomposition theorem)

Let $\eta_4(x_1, x_2, \cdots, x_n)$ be a homogeneous polynomial of degree 4 in $x_1, x_2, \cdots, x_n$ over $\mathbb{R}$. Let $H(\eta_4) = \left(\frac{\partial^2 \eta_4}{\partial x_i \partial x_j}\right)$, $k+1 \leq i, j \leq n$, be the Hessian matrix of η_4. Then $H(\eta_4)$ can not be decomposed as $\Delta(x)\Delta(x)^T$, where $\Delta(x) = (\beta_{ij})_{1 \leq i,j \leq n}$ is an anti-symmetric matrix with β_{ij} linear functions in x, unless η_4 and Δ are trivial, i.e., $H(\eta_4) = \Delta(x)\Delta(x)^T$ implies $\Delta = 0$ and $\eta_4 = 0$.

[11] Wu, Yau and Hu, *preprint*,1997

[12] Yau, Wu and Wong, *Mathematical Research Letters*,1999

Classification of maximal rank estimation algebra

Step 4. This final step was also done in 1997. Yau and Hu[13] used the full power of the quadratic structure theory developed by Chen and Yau (1997)[14] to prove that the matrix part (3) is with the constant coefficients.

[13] Yau and Hu, *preprint*,1997
[14] Chen and Yau, *SIAM J. Control Optim.*,1997

Classification of maximal rank estimation algebra

The above four steps complete the classification of FDEA of maximal rank. Therefore, Yau and his coworkers have proved the following theorem[15].

Theorem 7 (Complete classification)

Suppose that the state space of the filtering system is of dimension n. If E is the finite dimensional estimation algebra with maximal rank, then $f = (\alpha_1, \cdots, \alpha_n) + \nabla\phi$, where ϕ is a smooth function and $\alpha_i, 1 \leq i \leq n$, are affine functions and E is a real vector space of dimension $2n+2$ with basis given by
$1, x_1, \cdots, x_n, D_1, \cdots, D_n, L_0.$

As an immediate result, Mitter conjecture holds for maximal rank FDEA, which states any function in E is an affine function.

[15] Yau, *Internat. J. Control*, 2003

Progress of classification of non-maximal rank case

At the beginning of 20 century, classification of non-maximal rank estimation algebras becomes a very important and difficult problem.

- 2006: Classification of estimation algebra with state dimension 2. (Wu and Yau)[16];
- 2018: Linear structure of Ω and Mitter conjecture of state dimension 3, rank 2 case (Shi and Yau)[17] [18];
- 2020: Existence of novel finite dimensional filters (Jiao and Yau)[19].

[16] Wu and Yau, *SIAM J. Control Optim.*, 2006
[17] Shi and Yau, *SIAM J. Control Optim.*, 2017
[18] Shi and Yau, *Internat. J. Control*, 2020
[19] Jiao and Yau, *SIAM J. Control Optim.*, 2020

Classification of estimation algebra with state dimension 2

The most basic situation in non-maximal rank estimation algebra is that state dimension equal to 2. In work of Wu and Yau, general considerations and approaches toward the classification are proposed. Some structural results are obtained. The properties of Euler operators and the solution to an underdetermined PDE are extended.

Theorem 8 (Complete classification theorem, Wu & Yau)

Let state dimension $n = 2$. If E is finite-dimensional, then
(1) if h_i's are constants, $E = \{L_0\}$ or $E = \{L_0, 1\}$.
(2) otherwise, Ω-matrix has constant entries. h_i's must be affine in x_1 and x_2. E has dimension of either 4, 5, or 6.

Classification of estimation algebra with state dimension 3 and linear rank 2

Later, Shi and Yau study the structure of FDEA with state dimension 3 and rank 2 by using the theories of the Euler operator and underdetermined PDE. A powerful tool of their proof is technique of infinite sequence of differential operators.

Theorem 9 (Shi & Yau)

Let E be finite dimensional estimation algebras with state dimension 3 and rank 2. Then Ω has linear structure, i.e., entries of Ω are polynomial of degree 1.

Theorem 10 (Mitter conjecture, Shi & Yau)

The Mitter conjecture holds for state dimension 3, linear rank 2 case, that is any function in estimation algebra E is affine in x.

Novel finite dimensional filters

In 2020, Jiao and Yau solved construction of non-Yau type finite dimensional filter on arbitrary state dimension.

Theorem 11 (Jiao & Yau)

Nonlinear filtering system is given by,

$$\begin{cases} dx_1 = (x_1 + x_2 + \cdots + x_n + \gamma(x_1 + x_2 + \cdots + x_n))dt + dw_1, \\ dx_j = \sum_{i \neq j} x_i dt + dw_j, 2 \leq j \leq n \\ dy_k = \sqrt{\frac{k}{k+1}}(-\frac{1}{k}x_3 - \frac{1}{k}x_4 - \cdots - \frac{1}{k}x_{k+2} + x_{k+3})dt + dv_k, 1 \leq k \leq n-3 \\ dy_{n-2} = \frac{1}{\sqrt{n-2}}((n-2)x_2 - x_3 - \cdots - x_n)dt + dv_{n-2}, \end{cases} \tag{10}$$

where γ is a C^∞ function. Then in this filtering system, entries of Wong's Ω-matrix are not necessarily to be constants or polynomials. Dimension of estimation algebra is $2n-2$ and linear rank of estimation algebra is $n-2$.

Drawbacks of the Lie algebraic method

The basic idea of the Lie algebraic method is that solving the DMZ equation is transformed into solving a series of ordinary differential equations (ODEs), Kolmogorov equation, and some first-order linear partial differential equations (PDEs). However, the basis of the estimation algebra must be known in this method.
The direct method is the other approach to solve DMZ equation which works well especially for the Yau filtering system. Comparing with the Lie algebra method, we do not need to solve the basis of the estimation algebra, as well as integrate several first-order linear PDEs.

Works on direct method

- We studied the direct method w.r.t. time-invariant systems [20] [21] [22] [23]. Under some assumptions, solving the DMZ equation can be transformed into solving a series of ODEs and a Kolmogorov equation.
- We extended the direct method to time-varying systems [24] [25]. and in the latest work in 2019, the general direct method we proposed can treat nearly most general Yau filtering problems under natural assumptions.

[20] Hu and Yau, *IEEE Trans Aerosp Electron Syst.*, 2002.
[21] Yau and Yau, *IEEE CDC*, 1994.
[22] Yau and Hu, *IEEE Trans. Automat. Contr.*, 2001.
[23] Yau and Yau, *IEEE Trans Aerosp Electron Syst.*, 2004.
[24] Chen, Luo and Yau, *IEEE Trans Aerosp Electron Syst.*, 2017.
[25] Chen, Shi and Yau, *IEEE Trans. Automat. Contr.*, 2019.

Assumptions

We shall introduce the direct method for time-varying system (1) proposed in 2019 [26]. We assume that $G(x_t, t) = G(t)$ and define $\bar{G}(t) \triangleq G(t)Q(t)G^T(t)$.

Furthermore, we consider the time-varying Yau filtering system, i.e.,

$$f(x,t) = L(t)x + l(t) + \nabla_x \phi(t,x), \tag{11}$$

where $L(t) = (l_{ij}(t)), 1 \leq i,j \leq n, l^T(t) = (l_1(t), \cdots l_n(t))$ and $\phi(t,x)$ is a C^∞ function w.r.t. x on $\mathbb{R}^n$.

[26] Chen, Shi and Yau, *IEEE Trans. Automat. Contr.*, 2019.

Assumption 1

$\bar{G}(t) = G(t)Q(t)G^T(t)$ *is a positive definite matrix.*

Since $\bar{G}(t)$ is positive definite, then we can find a positive definite matrix $\Pi(t) > 0$ such that

$$\bar{G}(t) = \Pi(t)\Pi^T(t) \tag{12}$$

according to Cholesky decomposition.

Kolmogorov forward equation

Proposition 1 (DMZ equation → Kolmogorov equation)

For every $\tau_{i-1} \leq t \leq \tau_i$, *let*

$$\tilde{u}_i(\tau_i, x) = \exp\left(\sum_{j=1}^{m} y_j(\tau_{i-1})h_j(x)\right) u_i(\tau_i, x), \tag{13}$$

then $\tilde{u}_i$ *satisfies the Kolmogorov forward equation (KFE):*

$$
\begin{cases}
\dfrac{\partial \tilde{u}_i}{\partial t}(t,x) = & \dfrac{1}{2}\displaystyle\sum_{\iota,j=1}^{n} \bar{G}_{\iota j}(t)\dfrac{\partial^2 \tilde{u}_i}{\partial x_\iota \partial x_j}(t,x) - \sum_{\iota=1}^{n} f_\iota \dfrac{\partial \tilde{u}_i}{\partial x_\iota}(t,x) \\
& - \left(\displaystyle\sum_{\iota=1}^{n} \dfrac{\partial f_\iota}{\partial x_\iota}(t,x) + \dfrac{1}{2} h^T S^{-1} h \right) \tilde{u}_i(t,x), \\
\tilde{u}_1(0,x) = & \sigma_0(x), \\
\tilde{u}_i(\tau_{i-1},x) = & \exp\left[h^T(x,\tau_{i-1}) S^{-1}(\tau_{i-1})(y_{\tau_{i-1}} - y_{\tau_{i-2}}) \right] \\
& \cdot \tilde{u}_{i-1}(\tau_{i-1},x), \\
& i = 2,3,\cdots k.
\end{cases}
\tag{14}
$$

Transformation

Proposition 2

Under the Assumption 1, and let $\tilde{u}_i(t,x)$ *be the solution of* (14) *in* $[\tau_{i-1},\tau_i]$, $i=1,2,\cdots k$, $f(x,t)$ *satisfies (11). Let*

$$\tilde{u}_i(t,x)=e^{\bar{\phi}(t,x)}\tilde{\psi}_i(t,x), \tag{15}$$

where $\bar{\phi}(t,x)$*satisfies* $\nabla_x\bar{\phi}(t,x)=\bar{G}^{-1}(t)\nabla_x\phi(t,x)$ *and* $\bar{G}(t)$ *is defined in Assumption 1, then* $\tilde{\psi}_i(t,x)$ *satisfies the following equation:*

$$\begin{cases} \frac{\partial \tilde{\psi}_i}{\partial t}(t,x) = & \frac{1}{2}\sum_{\iota,j=1}^{n} \bar{G}_{\iota j}(t)\frac{\partial^2 \tilde{\psi}_i}{\partial x_\iota \partial x_j}(t,x) \\ & -(Lx+l)^T \nabla \tilde{\psi}_i(t,x) - \frac{1}{2}\bar{q}(t,x)\tilde{\psi}_i(t,x), \\ \tilde{\psi}_1(0,x) = & \sigma_0(x)e^{-\bar{\phi}(0,x)}, \\ \tilde{\psi}_i(\tau_{i-1},x) = & \exp\left[h^T(x,\tau_{i-1})S^{-1}(\tau_{i-1})(y_{\tau_{i-1}} - y_{\tau_{i-2}})\right] \\ & \cdot \tilde{\psi}_{i-1}(\tau_{i-1},x), \; i = 2,3,\cdots k, \end{cases} \tag{16}$$

where

$$\bar{q}(t,x) = -\sum_{i,j=1}^{n} \bar{G}_{ij}(t)\frac{\partial^2 \bar{\phi}}{\partial x_i \partial x_j}(t,x) + \nabla_x \bar{\phi}^T(t,x)\bar{G}(t)\nabla_x \bar{\phi}(t,x)$$
$$+ 2(Lx+l)^T \nabla_x \bar{\phi}(t,x) + 2\sum_{i=1}^{n} \frac{\partial^2 \phi(t,x)}{\partial^2 x_i^2} + 2\frac{\partial \bar{\phi}(t,x)}{\partial t}$$
$$+ \sum_{p,l=1}^{n} S_{pl}^{-1}(t) h_p(x,t) h_l(x,t) + 2tr(L). \tag{17}$$

Transformation

Theorem 12

Under the Assumption 1, and $\tilde{\psi}_i(t,x)$ is the solution of (16), let

$$\tilde{\psi}_i(t,x) = \psi_i(t,z), \tag{18}$$

where

$$\begin{aligned} z &= B(t)x, \\ B(t) &= \Pi^{-1}(t), \end{aligned} \tag{19}$$

and $\Pi(t)$ is defined in equation (12).

Then $\psi_i(t,z)$ is the solution of the following equation:

$$\begin{cases} \dfrac{\partial \psi_i}{\partial t}(t,x) = \dfrac{1}{2}\Delta\psi_i(t,x) - \dfrac{1}{2}\tilde{\bar{q}}(t,x)\psi_i(t,x) \\ \qquad\qquad - \left[\left(\dfrac{dB}{dt}B^{-1} + BLB^{-1}\right)x + Bl\right]^T \nabla\psi_i(t,x) \\ \psi_1(0,x) = \sigma_0(\Pi(0)x)\exp\left(-\bar{\phi}(0,\Pi(0)x)\right) \\ \psi_i(\tau_{i-1},x) = \exp\left[h^T(\Pi(\tau_{i-1})x,\tau_{i-1})S^{-1}(\tau_{i-1})\right. \\ \qquad\qquad \left.(y_{\tau_{i-1}} - y_{\tau_{i-2}})\right]\psi_{i-1}(\tau_{i-1},x), \\ \qquad\qquad i = 2,3,\cdots k. \end{cases} \tag{20}$$

where

$$\tilde{\bar{q}}(t,z) := \bar{q}(t,\Pi(t)z), \tag{21}$$

Assumption

Assumption 2

$\tilde{\bar{q}}(t,x)$ *in (21) is quadratic w.r.t.* x.

It follows naturally $\tilde{\bar{q}}(t,x)$ can be rewritten as

$$-\frac{1}{2}\tilde{\bar{q}}(t,x) = x^T\bar{Q}(t)x + p^T(t)x + r(t), \tag{22}$$

where $\bar{Q}(t)$ is a $n \times n$ symmetric matrix, $p(t)$ is a $n \times 1$ vector and $r(t)$ is a scalar.

Main theorem

Theorem 13

Under Assumption 1 and Assumption 2, we consider the following equation:

$$\begin{cases} \dfrac{\partial \psi_i}{\partial t}(t,x) = \dfrac{1}{2}\Delta\psi_i(t,x) - \dfrac{1}{2}\tilde{\tilde{q}}(t,x)\psi_i(t,x) \\ \qquad\qquad - \left[\left(\dfrac{dB}{dt}B^{-1} + BLB^{-1}\right)x + Bl\right]^T \nabla\psi_i(t,x) \\ \psi_i(\tau_{i-1},x) = \exp\left\{x^T A(\tau_{i-1})x + b^T(\tau_{i-1})x + c(\tau_{i-1})\right\}, \end{cases} \tag{23}$$

where $A(\tau_{i-1})$ *is a* $n \times n$ *symmetric matrix,* $b(\tau_{i-1})$ *is a* $n \times 1$ *column vector,* $x^T = (x_1, x_2, \dots, x_n)$ *is a row vector and* $c(\tau_{i-1})$*is a scalar. Then the solution of* (23) *is of the following form:*

$$\psi_i(t,x) = \exp\left\{x^T A(t)x + b^T(t)x + c(t)\right\}, \tag{24}$$

where $A(t)$ is a $n \times n$ matrix function w.r.t. t which is symmetric, $b(t)$ is a $n \times 1$ column vector function w.r.t. t and $c(t)$ is a scalar function w.r.t. t, and they satisfy the following ODEs:

$$\begin{aligned} \frac{dA(t)}{dt} &= 2A^2(t) - 2A(t)D(t) + \bar{Q}(t), \\ \frac{db^T(t)}{dt} &= 2b^T(t)A(t) - b^T(t)D(t) - 2d^T(t)A(t) + p^T(t), \\ \frac{dc(t)}{dt} &= trA(t) + \frac{1}{2}b^T b(t) - d^T(t)b(t) + r(t), \end{aligned} \tag{25}$$

where

$$D(t) = \frac{dB}{dt}B^{-1} + BLB^{-1}, d(t) = B(t)l(t). \tag{26}$$

Theorem 13 requires that the initial value $\psi_i(\tau_{i-1}, x)$ at every τ_{i-1} must be Gaussian. We first give a new way to do Gaussian approximation in algorithm 1, then non-Gaussian function can be approximated by the sum of several Gaussian functions and we can use Theorem 13 for every Gaussian function. The general direct method is summarized in Algorithm 2.

Gaussian approximation

Algorithm 1 Gaussian approximation

1: Let $f(x) = \phi(x)$ and the threshold $E = \alpha \cdot \max \phi(x)$, where α is a given small number.

2: Fitting the peaks of $f(x)$ which are larger than E with gaussian distributions. Suppose the sum of gaussian distributions in this step is $g(x)$.

3: Let $f_1(x) = f(x) - g(x)$. If $f_1(x)$ has no peaks whose values larger than E, then go to step 4. Otherwise, let $f(x) = f_1(x)$ and go to step 2.

4: Let $f_2(x) = -f_1(x)$. If $f_2(x)$ has no peaks which are larger than E, then done. Otherwise, let $f(x) = f_2(x)$ and go to step 2.

Algorithm 2 General direct method for time-varying system

1: **Initialization**: give $T, \Delta t, \sigma_0(x)$ and the parameter α in Algorithm 1. Let $k = \frac{T}{\Delta t}$, and $\{0 = \tau_0 < \tau_1 < \tau_2 < ... < \tau_k = T\}$.

2: **for** $i = 1 : k$ **do**

3: Using Algorithm 1 to get the Gaussian approximation $\psi_i(\tau_{i-1}, x) \approx \sum_{\iota=1}^{k(i)} \alpha_{i,\iota} \mathcal{N}(\mu_{i,\iota}, \sigma_{i,\iota})$.

4: For each Gaussian distribution $\mathcal{N}(\mu_{i,\iota}, \sigma_{i,\iota})$, suppose the solution of (23) with initial condition $\mathcal{N}(\mu_{i,\iota}, \sigma_{i,\iota})$ is $\psi_{i,\iota}(\tau_i, x)$. Solving (25), we obtain $\psi_{i,\iota}(\tau_i, x)$. Then $\psi_i(\tau_i, x) = \sum_{\iota=1}^{k(i)} \alpha_{i,\iota} \hat{u}_{i,\iota}(\tau_i, x)$.

5: Calculate $\psi_{i+1}(\tau_i, x)$ by $\psi_i(\tau_i, x)$ and (20).

6: Calculate $\tilde{\psi}_i(t_i, x), \tilde{u}_i(t_i, x)$ by (18),(15).

7: Calculate $u_i(t_i, x), \sigma(t_i, x)$.

8: Calculate conditional density function by normalization.

9: **end for**

Introduction

- In 2008 [1], Yau and Yau show that the DMZ equation (4) admits a unique nonnegative weak solution u which can be approximated by a solution u_R of the DMZ equation on the ball B_R with $u_R|_{\partial B_R} = 0$. The error of this approximation is bounded by a function of R which tends to zero as R goes to infinity. The solution u_R can in turn be approximated efficiently by an algorithm depending only on solving the observation-independent Kolmogorov equation on B_R.

[1] Yau and Yau, *SIAM J. Control Optim.*, 2008

- In 2013 [27], Luo and Yau extend the algorithm developed previously by Yau and Yau to the most general setting of nonlinear filterings, where the explicit time-dependence is in the drift term, observation term, and the variance of the noises could be a matrix of functions of both time and the states.

  There are some works investigating Hermite spectral method [28], proper orthogonal decomposition method [29] and Legendre spectral method [30], to numerically solve the forward Kolmogorov equation which help to solve the DMZ equation.

[27]Luo and Yau, *IEEE Trans. Automat. Contr.*, 2013a
[28]Luo and Yau, *IEEE Trans. Automat. Contr.*, 2013b
[29]Wang, Luo, Yau and Zhang, *IEEE Trans. Automat. Contr.*, 2020
[30]Dong, Luo and Yau, *IEEE Trans. Automat. Contr.*, 2013

Assumptions

In the work of Luo and Yau [31], the general case of (1) was considered:

Assumptions:

1. The operator L defined in (3) is uniform elliptic.
2. $||GQG^T||_\infty < \infty$, for all $(x,t) \in \mathbb{R}^n \times [0,T]$.
3. The initial density function $\sigma_0(x)$ decays fast enough. Namely, $\int_{\mathbb{R}^n} e^{\sqrt{1+|x|^2}}\sigma_0(x)dx < \infty$.

[31] Luo and Yau, *IEEE Trans. Automat. Contr.*, 2013a

Notations

Let us denote

$$D_w g := \left[\sum_{i=1}^{n} (GQG^T)_{ij} \frac{\partial g}{\partial x_i} \right]_{j=1}^{n}; \tag{27}$$

$$D_w^2 g := \sum_{i,j=1}^{n} (GQG^T)_{ij} \frac{\partial^2 g}{\partial x_i \partial x_j}; \tag{28}$$

$$K(x,t) := h^T(x,t) S^{-1}(t) y_t, \tag{29}$$

and

$$\begin{aligned} N(x,t) := & -\frac{\partial}{\partial t}\left(h^T S^{-1}\right) y_t - \frac{1}{2} D_w^2 K + \frac{1}{2} D_w K \cdot \nabla K \\ & - f \cdot \nabla K - \frac{1}{2}\left(h^T S^{-1} h\right). \end{aligned} \tag{30}$$

Step 1 for validation

Theorem 14 (u_R is a good approximation of u)

For any $T > 0$. Assume the following conditions are satisfied, for all $(x,t) \in \mathbb{R}^n \times [0,T]$:

1. $N(x,t) + \frac{3}{2}n \left|\left|GQG^T\right|\right|_\infty + |f - D_w K| \leq C,$ (31)
2. $e^{-\sqrt{1+|x|^2}} \left[14n \left|\left|GQG^T\right|\right|_\infty + 4\left|f - D_w K\right|\right] \leq \tilde{C},$ (32)

where C, $\tilde{C}$ are constants possibly depending on T. Let $R \gg 1$. Then $v := u - u_R \geq 0$ for all $(x,t) \in B_R \times [0,T]$ and

$$\int_{B_{\frac{R}{2}}} v(x,T) \leq Ce^{-\frac{9}{16}R} \int_{\mathbb{R}^n} e^{\sqrt{1+|x|^2}} \sigma_0(x), \tag{33}$$

where $C = C(T)$.

Step 2 for validation

Theorem 15 (u_R is well approximated by $u_{i,R}$, as $|\mathcal{P}_k| \to 0$)

Assume that

1. $N(x,t) \leq C,$ *(34)*
2. *There exists some $\alpha \in (0,1)$, such that*

$$|N(x,t) - N(x,t;\bar{t})| \leq \tilde{C}|t-\bar{t}|^{\alpha}, \quad (35)$$

for all $(x,t) \in B_R(0) \times [0,T]$, $\bar{t} \in [0,T]$, and $N(x,t;\bar{t})$ denotes the observation y_t contained in $N(x,t)$ be freezed at $y_{\bar{t}}$.
For any $0 \leq \tau \leq T$, let $\mathcal{P}_k = \{0 = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_k = \tau\}$ be a partition of $[0,\tau]$, where $\tau_i = \frac{i\tau}{k}$.

Then

$$u_R(x,\tau) = \lim_{k\to\infty} u_{k,R}(x,\tau),$$

in the L^1 sense in space and the following estimate holds:

$$\int_{B_R(0)} |u_R - u_{k,R}|(x,\tau) \leq C\frac{1}{k^\alpha}, \tag{36}$$

where $C = C(T, \int_R \sigma_0)$. The right-hand side of (36) tends to zero as $k \to \infty$.

Aim and assumptions

Aim of design: Two features have to be kept:

1. in real time: instantaneous feedback;
2. without memory: previous observation data are not necessary.

Assumptions:

- The time sequence of the observation is known apriopriliy, denote as $\mathcal{P}_k = \{0 = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_k = T\}$.
- The observation data $\{y_{\tau_i}\}_{i=0}^k$ are unknown until the on-line experiment runs.

Observation

Observation:

For each $\tau_{i-1} \leq t \leq \tau_i$, let

$$\tilde{u}_i(x,t) = \exp\left[h^T(x,t)S^{-1}(t)y_{\tau_{i-1}}\right]u_i(x,t), \tag{37}$$

then $\tilde{u}_i(x,t)$ satisfies the forward Kolmogorov equation (FKE)

$$\frac{\partial \tilde{u}_i}{\partial t}(x,t) = \left(L - \frac{1}{2}h^T S^{-1} h\right)\tilde{u}_i(x,t). \tag{38}$$

Advantage:

$(L - \frac{1}{2}h^T S^{-1} h)$ is independent of y_t. Hence, $\tilde{u}_i$ can be pre-computed.

Design of Yau-Yau algorithm

Off-line:

- Pick $\{\phi_n(x)\}_{n=0}^{\infty}$: a set of complete orthonormal base in $L^2(\mathbb{R})$.
- Compute the solution of FKE (38) at time τ_i with initial value $\phi_n(x)$ at τ_{i-1}, store each solution denoted as $\mathcal{U}(\tau_i, \tau_{i-1})\phi_n(x)$.

On-line:

- Update $\tilde{u}_{i+1}(x, \tau_i)$ by $\tilde{u}_{i+1}(x, \tau_i) = \exp\left[h^T(x, \tau_i)S^{-1}(\tau_i)(y_{\tau_i} - y_{\tau_{i-1}})\right]\tilde{u}_i(x, \tau_i)$.
- Projection: $\tilde{u}_i(x, \tau_{i-1}) = \sum_{n=0}^{\infty} \hat{c}_{i,n}\phi_n(x) \in L^2(\mathbb{R})$.
- Synchronize with off-line data: $\tilde{u}_i(x, \tau_i) = \sum_{n=0}^{\infty} \hat{c}_{i,n}\left[\mathcal{U}(\tau_i, \tau_{i-1})\phi_n(x)\right]$.

Experiment of almost linear filtering

We consider

$$\begin{cases} dx_t = dv_t \\ dy_t = x_t(1 + 0.25\cos x_t)dt + dw_t, \end{cases} \tag{39}$$

where x_t, $y_t \in \mathbb{R}$ and $E[dv_t^T dv_t] = E[dw_t^T dw_t] = 1$. Suppose the signal at the beginning is somewhere near 0 (or somewhere near 9).

The FKE from our model (39) is

$$u_t = \frac{1}{2}u_{xx} - \frac{1}{2}x^2(1 + \cos x)^2 u \tag{40}$$

Assume further that the initial distribution of x_0 is $u_0(x) = e^{\frac{-x^2}{2}}$ (or $u_0(x) = e^{\frac{-(x-9)^2}{2}}$).

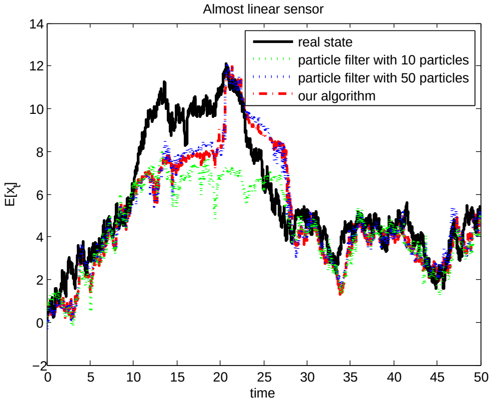


Figure 1: Almost linear filter is investigated with our algorithm and the particle filter with 10 and 50 particles. The total experimental time is $T = 50s$. And the update time is $\triangle t = 0.01$.

Comparison in efficiency: The CPU times of particle filter with 10 and 50 particles are 5.00s and 35.75s, respectively, while that of our algorithm is only 2.62s.

Experiment of cubic sensor

We consider

$$\begin{cases} dx_t = dv_t \\ dy_t = x_t^3 dt + dw_t, \end{cases} \tag{41}$$

where x_t, $y_t \in \mathbb{R}$ and $E[dv_t^T dv_t] = 1$, $E[dw_t^T dw_t] = 1$. Assume the initial state is somewhere near 0.

The KFE is

$$u_t = \frac{1}{2} u_{xx} - \frac{1}{2} x^6 u. \tag{42}$$

Furthermore, we assume the initial distribution is $u_0(x) = e^{-x^4/4}$.

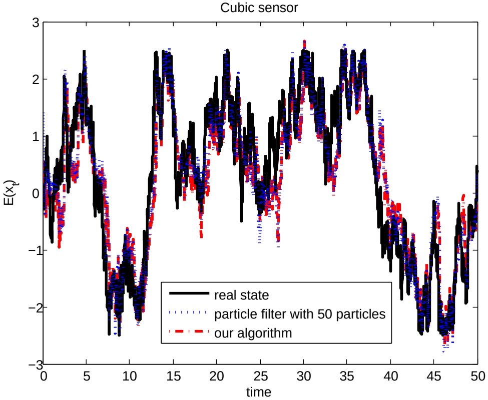
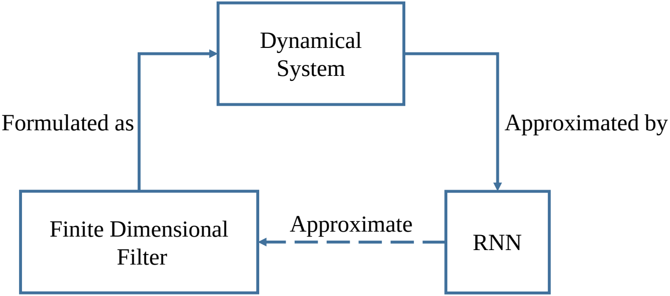


Figure 2: Cubic sensor for $T = 50$, with the time step $\triangle t = 0.01$s, by both particle filter and our algorithm.

Comparison in efficiency: The CPU time for our algorithm is 4.9s, while that for the particle filter is 37.17s. Our algorithm costs 0.001s to update, which is 10 times less than the update time.

Framework

In this work, we formulate the finite dimensional filter as the dynamical system with stochastic inputs[32]. Therefore, one natural idea is to approximate finite dimensional filters by RNN, i.e., we can solve filtering problems by RNN.



[32] Chen, Tao, Xu and Yau, *IEEE Trans. Neural Netw. Learn. Syst.*, 2022

Universal Approximation of Multilayer Feedforward Networks

Let $\Sigma^{r,N}(\kappa)$ be the class of functions

$$\{\bar{\zeta} = (\bar{\zeta}_1, \cdots, \bar{\zeta}_N)^{\mathrm{T}} : \mathbb{R}^r \to \mathbb{R}^N : \bar{\zeta}_l(x) = \sum_{j=1}^{q} \beta_{l,j}\kappa(A_j(x)),$$

$$x \in \mathbb{R}^r, \beta_{l,j} \in \mathbb{R}, A_j \in \mathbf{A}^r, 1 \le l \le N, q = 1, 2, \cdots\},$$

where $\kappa : \mathbb{R} \to [0,1]$ is the activation function and A_j is affine function. Apparently, $\bar{\zeta}$ represents the standard three-layered feedforward network with r input-neurons, q hidden-neurons and N output-neuron, which is shown in Fig. 3. It is well-known that this class of feedforward network functions are capable to approximate any continuous function over a compact set to any desired degree of accuracy[33].

[33] Hornik, Stinchcombe and White, *Neural Netw.* 1989

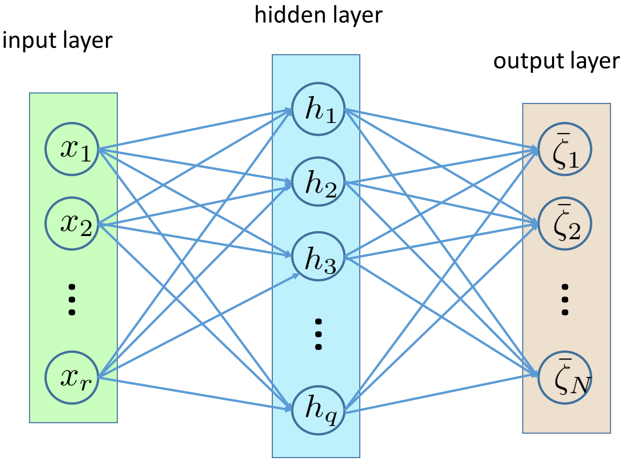


Figure 3: Three-layered feedforward network with r input-neurons, q hidden-neurons and N output-neuron.

Universal approximation of RNN with stochastic inputs

An open dynamical system in discrete time can be represented by the following equations:

$$\begin{cases} s_{k+1} = \eta(s_k, \alpha_{k+1}), & \text{state transition} \\ \beta_k = \xi(s_k), & \text{output equation} \end{cases} \tag{43}$$

where α_k is the stochastic external input, s_k is the state and β_k is the observable output for $\forall\ k \geq 1$.
Now we aim to approximate the open dynamical system (43) with stochastic inputs by a class of RNNs. We define the truncation operator $\mathcal{T}_K$ with level $K > 0$ as

$$\mathcal{T}_K(x_i) = \begin{cases} x_i & \text{if} \quad |x_i| \leq K \\ K \cdot \operatorname{sign}(x_i) & \text{otherwise}, \end{cases} \tag{44}$$

and $\mathcal{T}_K(x) := (\mathcal{T}_K(x_1), \cdots, \mathcal{T}_K(x_n))^{\mathrm{T}}$ for $x = (x_1, \cdots, x_n)^{\mathrm{T}}$.

Definition 16

For any squashing function κ, and $r_1, r_2, r_3 \in \mathbb{N}$, $RNN^{r_1,r_2,r_3}(\kappa)$ is a class of functions with the following state space model form:

$$\begin{cases} \tilde{s}_{k+1} = \tilde{\eta}(\tilde{s}_k, \alpha_{k+1}), \\ \tilde{\beta}_k = \tilde{\xi}(\tilde{s}_k), \end{cases} \tag{45}$$

where $\alpha_k \in \mathbb{R}^{r_1}$ is the input, $\tilde{s}_k \in \mathbb{R}^{r_2}$ is the hidden state, $\tilde{\beta}_k \in \mathbb{R}^{r_3}$ is the output, and

$$\tilde{\eta}(\tilde{s}, \alpha) = \bar{\eta}(\mathcal{T}_{K^s}\tilde{s}, \mathcal{T}_{K^\alpha}\alpha), \tag{46}$$

$$\tilde{\xi}(\tilde{s}) = \bar{\xi}(\mathcal{T}_{K^s}\tilde{s}), \tag{47}$$

in which $\bar{\eta} \in \Sigma^{r_1+r_2,r_2}(\kappa)$, $\bar{\xi} \in \Sigma^{r_2,r_3}(\kappa)$, K^s and K^α are two positive numbers which are the parameters of the RNN.

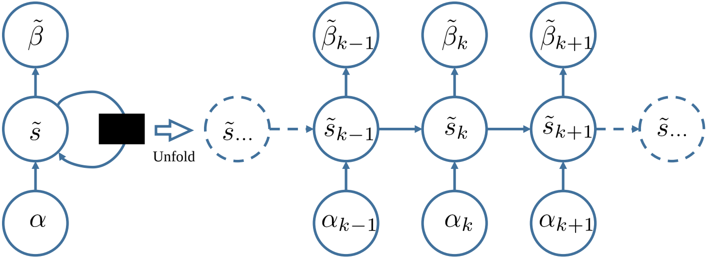


Figure 4: Recurrent neural networks with input α, hidden state $\tilde{s}$ and output $\tilde{\beta}$.

Theorem 17

Let $\eta(\cdot) : \mathbb{R}^{r_2} \times \mathbb{R}^{r_1} \to \mathbb{R}^{r_2}$ *and* $\xi(\cdot) : \mathbb{R}^{r_2} \to \mathbb{R}^{r_3}$ *be continuous, the external stochastic inputs* $\alpha_k \in \mathbb{R}^{r_1}$*, the inner state* $s_k \in \mathbb{R}^{r_2}$*, and the output* $\beta_k \in \mathbb{R}^{r_3}$*,* $k = 1, 2, \cdots$*. For any open dynamical system of the form* (43)*, if the following conditions hold:*

- $\{\alpha_k, k \geq 1\}$ *and* $\{s_k, k \geq 1\}$ *are uniformly integrable*[a]*;*
- *for* $\forall\ s, \bar{s} \in L^1(\Omega; \mathbb{R}^{r_2})$ *and* $\forall\ \alpha, \bar{\alpha} \in L^1(\Omega; \mathbb{R}^{r_1})$*,* $\|\eta(s, \alpha) - \eta(\bar{s}, \bar{\alpha})\|_1 \leq C_{\eta 1} \|s - \bar{s}\|_1 + C_{\eta 2} \|\alpha - \bar{\alpha}\|_1$*, and the Lipschitz constant* $C_{\eta 1}$ *satisfies* $|C_{\eta 1}| < 1$*;*
- *for* $\forall\ \epsilon > 0$*, there exists* $\delta > 0$*, such that for any* $s, \bar{s} \in L^1(\Omega; \mathbb{R}^{r_2})$ *satisfying* $\|s - \bar{s}\|_1 < \delta$*, we have* $\|\xi(s) - \xi(\bar{s})\|_1 < \epsilon$*,*

[a] $\{\alpha_k, k \geq 1\}$ is uniformly integrable means that $\lim_{M \to +\infty} \left(\sup_{k \geq 1} \mathbb{E}[|\alpha_k| \mathbb{I}_{|\alpha_k| > M}] = 0 \right)$.

then (43) can be approximated by the functions in $RNN^{r_1,r_2,r_3}(\kappa)$ with an arbitrary accuracy, i.e., for $\forall\ \varepsilon > 0$, there exist functions $\tilde{\eta}$ and $\tilde{\xi}$ of forms (46)-(47), which determine the RNN system (45) with the same input $\{\alpha_k, k \geq 1\}$ of (43), such that[a]

$$\begin{aligned} \overline{\lim_{k\to\infty}} \left\| s_k - \tilde{s}_k \right\|_1 < \varepsilon, \\ \overline{\lim_{k\to\infty}} \left\| \beta_k - \tilde{\beta}_k \right\|_1 < \varepsilon, \end{aligned} \tag{48}$$

where $\tilde{s}_k$ and $\tilde{\beta}_k$ are the state and output of the RNN system (45), respectively.

[a]The norm $\|\cdot\|_1 := \mathbb{E}[|\cdot|]$.

RNN based finite dimensional filters

The discrete time-invariant filtering system considered here is as follows:

$$\begin{cases} x_k = f(x_{k-1}) + g(x_{k-1})w_{k-1}, \\ y_k = h(x_k) + v_k. \end{cases} \tag{49}$$

For finite dimensional filters, we have the following evolution functions of the sufficient statistics $S_{k|k}$ and the optimal estimate $\mathbb{E}[x_k|Y_k]$:

$$\begin{cases} S_{k|k} = \Phi(S_{k-1|k-1}, y_k), \\ \mathbb{E}[x_k|Y_k] = \Gamma(S_{k|k}), \end{cases} \tag{50}$$

where Φ and Γ are some functions determined by the system (49) and the explicit forms may be unknown.

Using the universal approximation of RNN in Theorem 17, we can approximate open dynamical system (50) by a RNN system which is as follows:

$$\begin{cases} \tilde{S}_{k|k} = \tilde{\Phi}(\tilde{S}_{k-1|k-1}, y_k), \\ \hat{x}_{k|k} = \tilde{\Gamma}(\tilde{S}_{k|k}), \end{cases} \tag{51}$$

where $\tilde{S}_{k|k}$ and $\hat{x}_{k|k}$ are defined as the state and output of the RNN system (51), respectively.
Using the data $\{y_k, \mathbb{E}[x_k|Y_k]\}_{k\geq 0}$, we can train RNN system (51) such that $\mathbb{E}[x_k|Y_k]$ can be well approximated by the output $\hat{x}_{k|k}$, which can be regarded as the estimate of the state x_k of (49) based on observation history Y_k. And we call this filtering method as RNN filter (RNNF).

Naturally, we aim to minimize

$$L_0(\theta) := \frac{1}{K_1+1}\mathbb{E}\left[\sum_{k=0}^{K_1} |\hat{x}_{k|k} - \mathbb{E}[x_k|Y_k]|^2\right], \tag{52}$$

where $K_1 \in \mathbb{N}$ is the total time step in training, and θ represents all the trainable parameters in RNNF which determines $\hat{x}_{k|k}$.
Observing that

$$\begin{aligned}&\mathbb{E}\left[|x_k - \hat{x}_{k|k}|^2\right]\\ =&\mathbb{E}\left[|x_k - \mathbb{E}[x_k|Y_k]|^2\right] + \mathbb{E}\left[|\mathbb{E}[x_k|Y_k] - \hat{x}_{k|k}|^2\right],\end{aligned}$$

it follows that

$$\text{argmin}_\theta\, L_0(\theta) = \text{argmin}_\theta\, L(\theta), \tag{53}$$

where

$$L(\theta) := \frac{1}{K_1+1}\mathbb{E}\left[\sum_{k=0}^{K_1} |\hat{x}_{k|k} - x_k|^2\right]. \tag{54}$$

Therefore, instead of data $\{y_k, \mathbb{E}[x_k|Y_k]\}_{k\geq 0}$ where $\mathbb{E}[x_k|Y_k]$ cannot be obtained in most cases, we only need data $\{y_k, x_k\}_{k\geq 0}$ which can be easily generated from the system (49). We need to remark that this step is crucial since it allows us to get accessible data.

Theorem 18

Consider a discrete filtering system (49) with optimal FDF. We make the following assumptions:

- *the sufficient statistics* $\left\{S_{k|k}\right\}_{k\geq 0}$ *and the observations* $\{y_k\}_{k\geq 0}$ *are uniformly integrable;*
- *function* Φ *is Lipschitz, i.e., for any* $S, \bar{S} \in \mathbb{R}^{n_S}$ *and* $y, \bar{y} \in \mathbb{R}^m$*,*

$$\left\|\Phi(S, y) - \Phi(\bar{S}, \bar{y})\right\|_1 \leq C_{\Phi 1}\left\|S - \bar{S}\right\|_1 + C_{\Phi 2}\left\|y - \bar{y}\right\|_1, \tag{55}$$

  where n_S *is the dimension of* $S_{k|k}$*,* $C_{\Phi 1}$ *and* $C_{\Phi 2}$ *are Lipschitz constants, and* $C_{\Phi 1}$ *satisfies* $|C_{\Phi 1}| < 1$*;*
- *for* $\forall\ \epsilon > 0$*, there exists* $\delta > 0$*, such that for any* $s, \bar{s} \in L^1(\Omega; \mathbb{R}^{n_S})$ *satisfying* $\|s - \bar{s}\|_1 < \delta$*, we have* $\|\Gamma(s) - \Gamma(\bar{s})\|_1 < \epsilon$*.*

then for any $\varepsilon > 0$, there exists a RNNF (51), i.e., there exist $\tilde{\Phi}$ and $\tilde{\Gamma}$ represented by feedforward networks, i.e.,

$$\tilde{\Phi}(s, y) = \bar{\Phi}(\mathcal{T}_{K_1} s, \mathcal{T}_{K_2} y), \tag{56}$$

$$\tilde{\Gamma}(s) = \bar{\Gamma}(\mathcal{T}_{K_1} s), \tag{57}$$

such that

$$\varlimsup_{k\to\infty} \left\| S_{k|k} - \tilde{S}_{k|k} \right\|_1 < \varepsilon, \tag{58}$$

and

$$\varlimsup_{k\to\infty} \left\| \hat{x}_{k|k} - \mathbb{E}[x_k|Y_k] \right\|_1 < \varepsilon. \tag{59}$$

Summary

In this report, we introduced several works on filtering problems. The first two works, finite dimensional filter and Yau-Yau algorithm, are based on DMZ equation and can be used to obtain the conditional density function of the state. Lie algebraic method and direct method are for systems with finite dimensional filters, and the former is more theoretical and we only need to solve a series of ODEs in the latter. Yau-Yau algorithm can be used to deal with the general nonlinear filtering problems and we only need to solve a Kolmogorov equation off-line. More recently, we also investigate to use deep learning in filtering problems, and this result has also been included in this report.

THANKS!