



# RBFNN-based Minimum Entropy Filtering for a Class of Stochastic Nonlinear Systems

DOI:

[10.1109/TAC.2019.2914257](https://doi.org/10.1109/TAC.2019.2914257)

## Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

## Citation for published version (APA):

Yin, X., Zhang, Q., Wang, H., & Ding, Z. (2019). RBFNN-based Minimum Entropy Filtering for a Class of Stochastic Nonlinear Systems. *IEEE Transactions on Automatic Control*.  
<https://doi.org/10.1109/TAC.2019.2914257>

## Published in:

IEEE Transactions on Automatic Control

## Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

## General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [uml.scholarlycommunications@manchester.ac.uk](mailto:uml.scholarlycommunications@manchester.ac.uk) providing relevant details, so we can investigate your claim.



# RBFNN-based Minimum Entropy Filtering for a Class of Stochastic Nonlinear Systems

Xin Yin, *Member, IEEE*, Qichun Zhang, *Member, IEEE*, Hong Wang, *Senior Member, IEEE*  
and Zhengtao Ding, *Senior Member, IEEE*

**Abstract**—This paper presents a novel minimum entropy filter design for a class of stochastic nonlinear systems which are subjected to non-Gaussian noises. Motivated by stochastic distribution control, an output entropy model is developed using radial basis function neural network (RBFNN) while the parameters of the model can be identified by the collected data. Based upon the presented model, the filtering problem has been investigated while the system dynamics have been represented. As the model output is the entropy of the estimation error, the optimal nonlinear filter is obtained based on the Lyapunov design which makes the model output minimum. Moreover, the entropy assignment problem has been discussed as an extension of the presented approach. To verify the presented design procedure, a numerical example is given which illustrates the effectiveness of the presented algorithm. The contributions of this paper can be summarized as 1) an output entropy model is presented using neural network; 2) a nonlinear filter design algorithm is developed as the main result and 3) a solution of entropy assignment problem is obtained which is an extension of the presented framework.

**Index Terms**—Minimum entropy filtering, stochastic nonlinear systems, non-Gaussian distribution, radial basis function neural network

## I. INTRODUCTION

SINCE Kalman filter was proposed as a linear optimal observer design [1], [2], a lot of extensions of Kalman filter have been developed, such as the extended Kalman filter (EKF) [3] and the unscented Kalman filter (UKF) [4]. However, most of these results focus on the Gaussian stochastic systems. Due to the fact that non-Gaussian noises widely exist in practice, the stochastic non-Gaussian distribution systems became a significant research topic for filtering. For example, the stochastic distribution of estimation error is investigated in [5].

Based on the B-spline approximation, the decoupled model has been presented to avoid of using the partial differential equation analysis [6]. This model provides a solution to the tracking problem of the output probability density function (PDF) [7]. Moreover, trying to make the distribution as sharp as possible in practice which leads to the minimum entropy control for the non-Gaussian distribution [8]. On the other

hand, the minimum entropy filtering algorithms have also been investigated [9], [10]. These results assume that the distributions of the random noises are known and the recursive formulation of probability density function leads to large workloads of computing. Meanwhile, the data driven minimum entropy filter via kernel density estimation (KDE) [11] was proposed in [12]–[14]. Basically, it is very difficult to implement the filter in real-time for the stochastic nonlinear system with unknown distribution of noises since the probability density function is difficult to obtain in real-time using kernel density estimation [11]. Another approach for nonlinear non-Gaussian filtering [15] is particle filtering for which is difficult to analyze the convergence of the estimation error. Therefore, it is very significant to develop a new filtering algorithm with theoretical analysis which can be implemented in real-time for practical applications.

To achieve the objective, a nonlinear filter is designed for a class of stochastic nonlinear systems while the nonlinear dynamics of the estimation errors have been transformed to the states of the stochastic distribution model [16]. In addition, the entropy of the estimation error can be described by the extended stochastic distribution model using RBF neural network. In other words, the non-linearity can be represented by a neural network model, and then the real-time implementation can be guaranteed once the model has been trained sufficiently. Notice that any type of neural network is available to describe the output entropy which leads to a potential framework for entropy control problem.

Following this approach, the structure of the filter is given firstly which leads to the dynamics of the estimation error. Then, the dynamics can be represented by the RBFNN-based entropy model and the filtering can be given using Lyapunov design method. In other words, the designed filtering signal can be used to minimize the entropy of estimation error while the convergence is also analyzed in mean-square sense. The simulation results are compared with the standard EKF design and the better performances indicate the effectiveness of the proposed filtering algorithm.

## II. FORMULATION

Consider the following discrete-time stochastic nonlinear system,

$$\begin{aligned} x_{k+1} &= f(x_k, u_k) + w_k \\ y_k &= Cx_k + \bar{v}_k \end{aligned} \quad (1)$$

where  $y$  and  $u$  denote the system output and input, respectively. Suppose that the output is bounded  $y \in [a, b]$ ,  $f(\cdot)$  is

X. Yin and Z. Ding are with the Sino-British Joint Lab, the School of Electrical and Electronic Engineering, University of Manchester, Manchester, M13 9PL, United Kingdom. e-mail: (see xin.yin@manchester.ac.uk; zhengtao.ding@manchester.ac.uk).

Q. Zhang is with the School of Engineering and Sustainable Development, De Montfort University, Leicester, LE1 9BH, United Kingdom. e-mail: (qichun.zhang@dmu.ac.uk).

H. Wang is with Pacific Northwest National Laboratory, Richland, WA 99352, United States, e-mail: (hong.wang@pnl.gov).

Manuscript received April 19, 2005; revised August 26, 2015.

a known general continuous nonlinear function,  $C$  is known constant matrix, and  $w$  and  $\bar{v}$  stand for the random noise with zero mean value.

*Assumption 1:* The nonlinear function  $f(\cdot)$  satisfies Lipschitz condition which implies that the following inequality holds.

$$|f(x_i, u) - f(x_j, u)| \leq M |x_i - x_j|, \forall i, j \in \mathbb{Z}^+ \quad (2)$$

where  $M$  is a real positive constant.

Based on the system model, the structure of the filter can be obtained as follows:

$$\hat{x}_{k+1} = f(\hat{x}_k, u_k) + g_k \quad (3)$$

where  $\hat{x}$  and  $\hat{y}$  denote the estimated state  $x$  and the estimated output  $y$  while the dynamic of estimation error  $\tilde{x}$  is given as follows:

$$\tilde{x}_{k+1} = f(x_k, u_k) - f(\hat{x}_k, u_k) - g_k + w_k \quad (4)$$

Similarly the output estimation error can be formulated as  $\tilde{y}_k = C\tilde{x}_k + \bar{v}_k$ . In this paper,  $g_k$  will be designed by  $\tilde{y}$  which forms the design objective of this paper. In particular, the filter design would be completed if the entropy of the estimation error  $\tilde{x}$  is minimized which implies that the randomness of the estimation is minimized. Due to the fact that  $\tilde{x}$  is unmeasurable, the entropy of output estimation error  $\tilde{y}$  can be optimized equivalently to attenuate the randomness of the estimation. Therefore, Rényi entropy is considered in this paper,

$$H_{\alpha, k}(\tilde{y}) = \frac{1}{1 - \alpha} \log \int_a^b \gamma_k^\alpha(\tilde{y}_k, u_k) d\tilde{y} \quad (5)$$

where  $\alpha \geq 0$  is the order of the entropy,  $\gamma_k(\cdot)$  denotes the probability density function of the estimation error for sampling time instance  $k$ .

In summary, the design objective is to find nonlinear function  $g_k$  to minimize the entropy  $H_{\alpha, k}(\tilde{y})$  for stochastic nonlinear system (1).

### III. RBFNN-BASED ENTROPY MODEL

Motivated by B-spline stochastic distribution model [7], the nonlinear function  $g_k$  can be obtained similar to the controller design for the stochastic distribution control. Therefore, a novel RBFNN-based entropy model is presented in this section while the dynamic relationship between  $H_{\alpha, k}(\tilde{y})$  and  $g_k$  is established. Normally, the filter design is based on the optimization of state estimation error  $\tilde{x}$ . However the output estimation error  $\tilde{y}$  is also available to use if the prediction of the output entropy can be obtained. In addition, the random noise  $\bar{v}_k$  is also taken into account if the relationship between  $H_{\alpha, k}(\tilde{y})$  and  $g_k$  is described.

#### A. The Structure of the Model

Suppose that the probability density function of the estimation error  $\tilde{y}$  can be represented by the RBF neural network with weighing vector  $V_k$ . Thus we recall the following model[6]

$$V_{k+1} = AV_k + Bg_k$$

$$\sqrt{\gamma_k(\tilde{y}_k, g_k)} = \frac{G}{\sqrt{V_k^T E V_k}} V_k + e_\gamma \quad (6)$$

where  $V = [v_1, v_2, \dots, v_m]^T$  is the weight vector and  $G = [G_1(\tilde{y}), G_2(\tilde{y}), \dots, G_m(\tilde{y})]$  while  $G_i(\tilde{y})$  and  $e_\gamma$  stand for the pre-specified Gaussian basis functions for the approximation of  $\gamma(\tilde{y}, g)$  and the approximation residual.  $E = \int_a^b G^T(\tilde{y}) G(\tilde{y}) d\tilde{y}$ . Matrices  $A$  and  $B$  are the coefficient of the weight update equation.

Notice that the entropy (5) can be further expressed by the following formula if the order of the entropy is selected as  $\alpha = 1/2$ , then we have

$$H_k = 2 \log \int_a^b \sqrt{\gamma_k(\tilde{y}_k, g_k)} d\tilde{y} \quad (7)$$

Combining the model (6) and the definition of the entropy (7), the following equation can be obtained.

$$H_k = 2 \log \left( \int_a^b \frac{G}{\sqrt{V_k^T E V_k}} V_k d\tilde{y} + \int_a^b e_\gamma d\tilde{y} \right)$$

$$= 2 \log \left( \sum_{i=1}^m \frac{v_i}{\sqrt{V_k^T E V_k}} \int_a^b G_i(\tilde{y}) d\tilde{y} + \int_a^b e_\gamma d\tilde{y} \right) \quad (8)$$

According to the neural network approximation theory [17], there exists a positive integer  $m$  such that the error  $e_\gamma < \epsilon$  for any pre-specified  $\epsilon > 0$ . Moreover, integral value of the pre-specified basis function  $G_i(\tilde{y})$  is equal to 1. As a result, Eq.(8) can be rewritten by

$$H_k = 2 \log \left( \frac{\bar{G}}{\sqrt{V_k^T E V_k}} V_k \right) \quad (9)$$

where  $\bar{G} = [1, 1, \dots, 1] \in \mathbb{R}^m$ .

To simplify the expression, we can further denote the weight vector as follows:

$$\bar{V}_k = \frac{V_k}{\sqrt{V_k^T E V_k}} \quad (10)$$

Based on this transformation, the output entropy model of the estimation error is given by

$$\bar{V}_{k+1} = \bar{A}\bar{V}_k + \bar{B}g_k$$

$$H_k = 2 \log(\bar{G}\bar{V}_k) \quad (11)$$

where matrices  $\bar{A}$  and  $\bar{B}$  are the coefficient of the transformed weight update equation for  $\bar{V}$ .

## B. Parametric Identification

In order to facilitate the system analysis, the elements of the coefficient matrices  $\bar{A}$  and  $\bar{B}$  should be identified. Thus we define the following function.

$$r_k = e^{\frac{1}{2}H_k} \quad (12)$$

Transforming the presented model, the input-output format of the system can be obtained.

$$r_k = \bar{G}\bar{V}_k = \bar{G}(I - z^{-1}\bar{A})^{-1}\bar{B}g_{k-1} \quad (13)$$

with further expansion

$$r_k = \sum_{i=1}^m a_i r_{k-i} + \sum_{i=1}^{m-1} \sum_{j=1}^m d_{ij} g_{k-i} \quad (14)$$

Notice that the expression above can be further rewritten by the following format,

$$r_k = \theta^T \Phi_k \quad (15)$$

where

$$\theta = [a_1, \dots, a_m, d_{11}, \dots, d_{1m}, \dots, d_{(m-1),1}, \dots, d_{(m-1),m}]^T \quad (16)$$

$$\Phi_k = [r_{k-i}, \dots, r_{k-m}, g_{k-1}, \dots, g_{k-1}, \dots, g_{k-m}, \dots, g_{k-m}]^T \quad (17)$$

Based on Eq.(15), the unknown parametric vector  $\theta$  can be estimated using recursive least square (RLS) algorithm [18]:

$$\theta(i+1) = \theta(i) + \frac{P(i-1)\Phi_k(\tilde{y}_k, g_k)\varepsilon(i)}{1 + \Phi_k^T(\tilde{y}_k, g_k)P(i-1)\Phi_k(\tilde{y}_k, g_k)} \quad (18)$$

$$\varepsilon(i) = r_k - \theta^T(i)\Phi_k(\tilde{y}_k, g_k) \quad (19)$$

$$P(i) = \left( I - \frac{P(i-1)\Phi_k(\tilde{y}_k, g_k)}{1 + \Phi_k^T(\tilde{y}_k, g_k)P(i-1)\Phi_k(\tilde{y}_k, g_k)} \right) \times P(i-1) \quad (20)$$

Since the RBFNN-based approximation with weights can be given for probability density function as follows, the weights  $\bar{V}_k$  for any sampling instance  $k$  can be estimated.

$$\sqrt{\gamma_k(\tilde{y}_k, g_k)} = G\bar{V}_k + e_\gamma \quad (21)$$

which leads to

$$\bar{V}_k = (G^T G)^{-1} G^T \sqrt{\gamma_k(\tilde{y}_k, g_k)} \quad (22)$$

while the Gaussian basis functions are orthogonal then  $(G^T G)^{-1}$  exists.

Moreover, the probability density function  $\gamma_k$  and entropy  $H_k$  can be approximated using the kernel density estimation based on the online/offline training with collected data set. Then the initial value of the weights can be obtained and the weights can also be adjusted for any sampling instance  $k$  if it is necessary. Finally, once the parametric vector and initial weights are estimated, the matrices  $\bar{A}$  and  $\bar{B}$  can be obtained which completes the setup of the presented model.

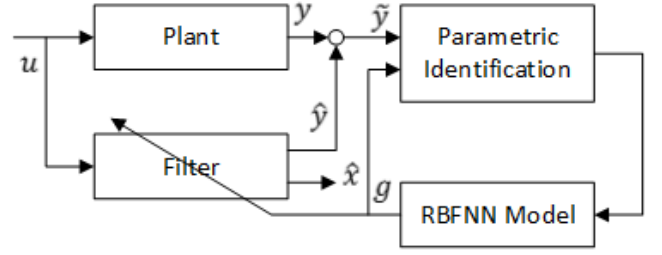


Fig. 1. The structure of the presented nonlinear filter.

In particular, the probability density function  $\gamma_k$  and entropy  $H_k$  can be estimated as follows once the bandwidth  $n_k$  is pre-specified.

$$\gamma_k = \frac{1}{n_k} \sum_{i=k-n_k}^k G(\eta - \tilde{y}_i) \quad (23)$$

where  $\eta$  denotes the random variable of  $\tilde{y}$ . Based on the estimated probability density function, the square-root of PDF and the entropy can be further estimated using numerical approach.

## IV. NONLINEAR FILTER DESIGN

Based upon the presented model (11), it is shown that the filtering problem has been transformed to an optimal control form. Particularly, the structure of the filter design can be described by Fig.IV.

To achieve the design objective, the performance criteria can be considered which consists of the entropy and the filtering adjustment signal with weights. Without loss of generality, the following format can be adopted to simplify the calculation,

$$J = \frac{1}{8}H_k^2 + \frac{\lambda}{2}g_k^2 \quad (24)$$

where  $\lambda$  denotes a pre-specified positive real number.

To minimise this performance criterion, the first order derivative is calculated which should be equal to 0.

$$\frac{\partial J}{\partial g} = 0 \quad (25)$$

which leads to

$$\begin{aligned} \frac{\partial J}{\partial g} &= \log(\bar{G}V_k) \frac{\partial \log(\bar{G}V_k)}{\partial g} + \lambda g_k \\ &= \frac{\log(\bar{G}V_k)}{\bar{G}V_k} \bar{G} \frac{\partial V_k}{\partial g} + \lambda g_k \\ &= \frac{\log(\bar{G}V_k)}{\bar{G}V_k} \bar{G} \frac{\partial V_k}{\partial t} / \frac{\partial g}{\partial t} + \lambda g_k = 0 \end{aligned} \quad (26)$$

Furthermore, we have

$$Q(\tilde{y}) \frac{\partial V_k}{\partial t} + \lambda g_k \frac{\partial g}{\partial t} = 0 \quad (27)$$

where

$$Q(\tilde{y}) = \frac{\log(\bar{G}V_k)}{\bar{G}V_k} \bar{G} \in \mathbb{R}^m \quad (28)$$

Thus the nonlinear filter can be designed and the discrete-time formula can be expressed using the backward difference with the sampling time  $\Delta t$  to replace the derivative.

$$\frac{\partial V_k}{\partial t} \approx \frac{V_k - V_{k-1}}{\Delta t}, \frac{\partial g}{\partial t} \approx \frac{g_k - g_{k-1}}{\Delta t} \quad (29)$$

which leads to

$$\lambda g_k^2 - g_k g_{k-1} + Q(\tilde{y})(V_k - V_{k-1}) = 0 \quad (30)$$

Solving this equation, the filter design can be completed by the following formula.

$$g_k = \frac{g_{k-1} \pm \sqrt{g_{k-1}^2 - 4\lambda Q(\tilde{y})(V_k - V_{k-1})}}{2\lambda} \quad (31)$$

where the sign in the formula can be determined by substituting  $g_k$  with the proper  $\lambda$  into following condition

$$\frac{\partial^2 J}{\partial g^2} > 0 \quad (32)$$

Moreover, the sign in Eq.(29) can be further determined by the following convergence analysis in next section.

Note that the performance criterion can also be considered as a Lyapunov function candidate. From Eq.(24), it shows that

$$\frac{\partial J}{\partial t} = \frac{\partial J}{\partial g} \frac{\partial g}{\partial t} = 0 \quad (33)$$

As we analysed that  $\frac{\partial J}{\partial g} = 0$  using the design function  $g$ , we claim that the filtering algorithm can be used to minimize the entropy of the estimation error.

As a summary, the pseudo-code of the presented filtering algorithm is given in this section.

---

#### Algorithm 1 Nonlinear Filtering

---

**Input:**  $\hat{x}_0, g_0, \bar{A}(0), \bar{B}(0), \lambda$

**Output:**  $\hat{y}, \hat{x}$

Measure system output  $y$  and initialization

**while**  $k > 0$  **do**

Update  $\hat{y}, \hat{x} \leftarrow$  Eq.(1) and Eq.(3)

**if**  $k <$  Polynomial order in Eq.(14) **then**

Off-line training based on pre-collected data

**else**

$a_i, d_{ij} \leftarrow$  RLS, Eq. (18-20)

**end if**

$\bar{A}(k), \bar{B}(k) \leftarrow$  Re-arrange  $a_i, d_{ij}$

$g_k^\pm \leftarrow$  Eq.(31)

$J_k^\pm \leftarrow$  Eq.(24)

**if**  $J_k^+ < J_k^-$  **then**

$g_k \leftarrow g_k^+$

**else**

$g_k \leftarrow g_k^-$

**end if**

$k \leftarrow k + 1$

**end while**

---

## V. CONVERGENCE ANALYSIS

Since the output estimation error is minimised with the presented filtering algorithm, the convergence of the estimation error  $\tilde{x}$  can be further analysed in mean square sense.

The dynamics of the estimation error Eq.(4) results in

$$E(\tilde{x}_{k+1}^2) = E(\tilde{f}_k^2 - 2\tilde{f}_k g_k + g_k^2) \quad (34)$$

where  $E(\cdot)$  denotes the mean-value operation with  $\tilde{f}_k = f(x_k, u_k) - f(\hat{x}_k, u_k)$ .

Similarly, we have

$$E(\tilde{x}_k^2) = E(\tilde{f}_{k-1}^2 - 2\tilde{f}_{k-1} g_{k-1} + g_{k-1}^2) \quad (35)$$

which leads to

$$E(\tilde{x}_{k+1}^2) - E(\tilde{x}_k^2) \leq E\left(\left(\tilde{f}_k - g_k\right)^2 + 2\tilde{f}_{k-1} g_{k-1}\right) \quad (36)$$

Notice that  $2\tilde{f}_{k-1} g_{k-1} = -\left(\tilde{f}_{k-1} - g_{k-1}\right)^2 + \tilde{f}_{k-1}^2 + g_{k-1}^2$ , and then Eq.(36) can be further expressed by

$$E(\tilde{x}_{k+1}^2) - E(\tilde{x}_k^2) \leq E\left(\left(\tilde{f}_k - g_k\right)^2 - \left(\tilde{f}_{k-1} - g_{k-1}\right)^2 + \tilde{f}_{k-1}^2 + g_{k-1}^2\right) \quad (37)$$

As there always exist real positive numbers  $\varepsilon_1, \varepsilon_2$  and  $\varepsilon_3$ , the following inequalities can be satisfied.

$$\left(\tilde{f}_k - g_k\right)^2 \leq \frac{1}{\varepsilon_1} \tilde{f}_k^2 + \varepsilon_2 g_k^2 \quad (38)$$

$$\left(\tilde{f}_{k-1} - g_{k-1}\right)^2 \geq \frac{1}{\varepsilon_1} \tilde{f}_{k-1}^2 - \varepsilon_3 g_{k-1}^2 \quad (39)$$

Based upon Assumption 1, there exists a positive constant  $M_k$  for each sampling instance  $k$  to satisfy inequality (2). Furthermore, it leads to that there always exists a positive constant  $N < \max(M_k^2)$  such that

$$\left(\tilde{f}_k - g_k\right)^2 \leq \frac{N}{\varepsilon_1} \tilde{x}_k^2 + \varepsilon_2 g_k^2 \quad (40)$$

$$\left(\tilde{f}_{k-1} - g_{k-1}\right)^2 \geq \frac{N}{\varepsilon_1} \tilde{x}_{k-1}^2 - \varepsilon_3 g_{k-1}^2 \quad (41)$$

Thus, Eq.(36) is transformed to

$$E(\tilde{x}_{k+1}^2) - E(\tilde{x}_k^2) \leq \frac{N}{\varepsilon_1} (E(\tilde{x}_k^2) - E(\tilde{x}_{k-1}^2)) + E\left((1 + \varepsilon_3) g_{k-1}^2 + \tilde{f}_{k-1}^2 + \varepsilon_2 g_k^2\right) \quad (42)$$

Note that Eq.(31) can be represented as follows:

$$g_k = g_{k-1} - \Delta_{k-1} \quad (43)$$

where

$$\Delta_{k-1} = \frac{(2\lambda - 1) g_{k-1} \pm \sqrt{g_{k-1}^2 - 4\lambda Q(\tilde{y})(V_k - V_{k-1})}}{2\lambda} \quad (44)$$

In addition, we have

$$\begin{aligned} \varepsilon_2 g_k^2 + (1 + \varepsilon_3) g_{k-1}^2 \\ = \varepsilon_2 (g_{k-1}^2 - 2g_{k-1}\Delta_{k-1} + \Delta_{k-1}^2) + (1 + \varepsilon_3) g_{k-1}^2 \\ = (1 + \varepsilon_2 + \varepsilon_3) g_{k-1}^2 - \varepsilon_2 (2g_{k-1}\Delta_{k-1} - \Delta_{k-1}^2) \end{aligned} \quad (45)$$

which results in the following inequality.

$$\begin{aligned} E(\tilde{x}_{k+1}^2) - E(\tilde{x}_k^2) \leq \frac{N}{\varepsilon_1} (E(\tilde{x}_k^2) - E(\tilde{x}_{k-1}^2)) + E(\tilde{f}_{k-1}^2) \\ + (1 + \varepsilon_2 + \varepsilon_3) g_{k-1}^2 - \varepsilon_2 (2g_{k-1}\Delta_{k-1} - \Delta_{k-1}^2) \end{aligned} \quad (46)$$

Since we can always find suitable positive constants  $\varepsilon_1$  and  $N$  make the following inequality holds,

$$0 < \frac{N}{\varepsilon_1} < 1 \quad (47)$$

thus the estimation error is convergent if the following condition is satisfied.

$$\tilde{f}_{k-1}^2 + (1 + \varepsilon_2 + \varepsilon_3) g_{k-1}^2 + \varepsilon_2 \Delta_{k-1}^2 - 2\varepsilon_2 g_{k-1} \Delta_{k-1} \leq 0 \quad (48)$$

To meet Eq.(32),  $\lambda$  can be chosen as  $\lambda \geq \frac{1}{2}$ , and then Eq.(44) shows that the plus sign can be used if  $g_{k-1} > 0$ . Similarly, the minus sign can be used if  $g_{k-1} < 0$  such that

$$\varepsilon_2 g_{k-1} \Delta_{k-1} > 0 \quad (49)$$

Since  $\varepsilon_2$  can be pre-specified as a large positive constant to satisfy the inequality (40), the condition (48) is implementable.

Finally, it can be represented as

$$|E(\tilde{x}_{k+1}^2) - E(\tilde{x}_k^2)| \leq \frac{N}{\varepsilon_1} |E(\tilde{x}_k^2) - E(\tilde{x}_{k-1}^2)| \quad (50)$$

It has been shown that the estimation error for each state of the presented stochastic system (1) is convergent with the filter (31), if there exist  $\varepsilon_1$ ,  $\varepsilon_2$ ,  $\varepsilon_3$  and  $\lambda$  meet the inequalities (47) and (48) for each time instance  $k$ . In addition, the sign in Eq.(31) can be chosen following the analysis (48) and (49).

## VI. A NUMERICAL EXAMPLE

To verify the presented filtering algorithm, the twin-level tank system is considered as a practical experiment while the structure is shown in Fig.2.

Following the system instruction in [19], the system model can be obtained as follows.

$$\begin{aligned} x_{1,k+1} &= -\frac{h}{A_1} (c_1 + k_1 \sqrt{x_{1,k} - k_0} \sqrt{x_{2,k} - x_{1,k}}) \\ &\quad + x_{1,k} + w_{1,k} \\ x_{2,k+1} &= \frac{h}{A_2} (k_4 u_{2,k} - c_2 - k_0 \sqrt{x_{2,k} - x_{1,k}}) \\ &\quad + x_{2,k} + w_{2,k} \\ y_k &= x_{1,k} + v_k \end{aligned} \quad (51)$$

where  $x_1$  and  $x_2$  stand for the levels of tank 1 and tank 2,  $A_1$  and  $A_2$  are the cross-sectional area,  $c_1$  and  $c_2$  are constant parameters of the valve and pump,  $k_0$ ,  $k_1$  and  $k_4$  denote the ratio of the valves, and  $h$  is the sampling time.

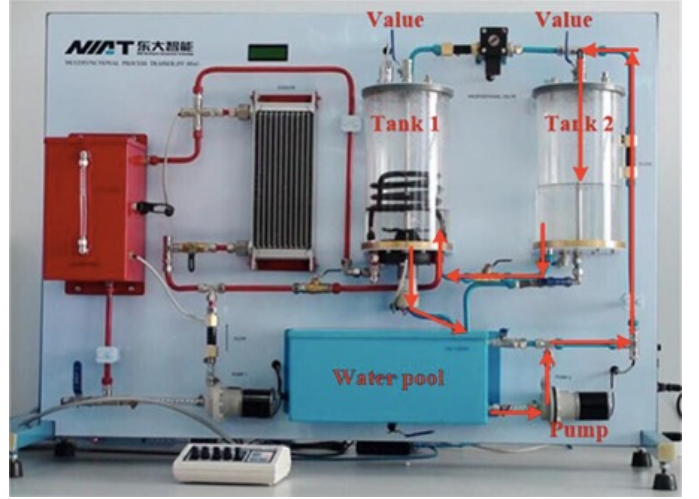


Fig. 2. Structure of the twin-tank level system.

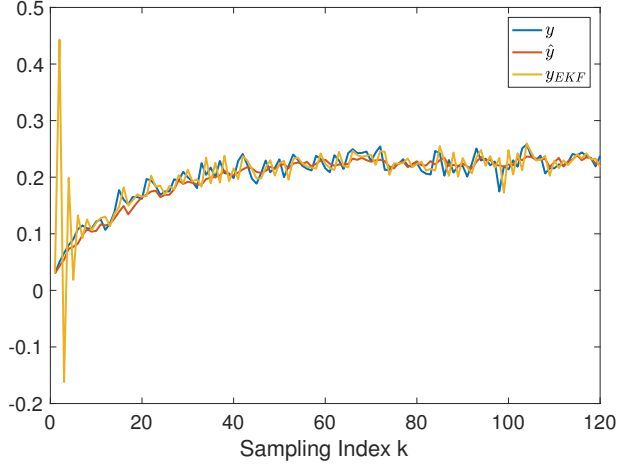
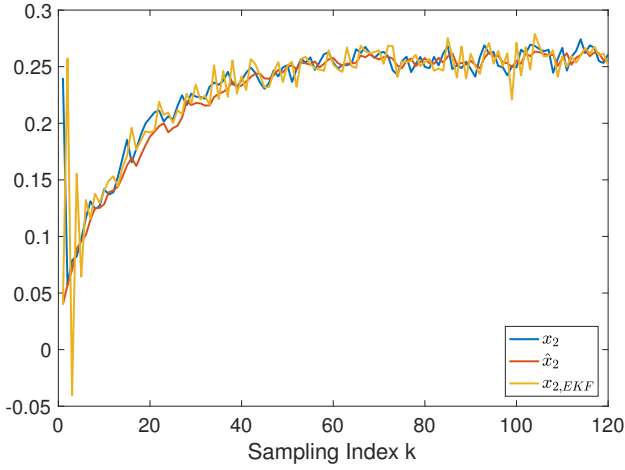
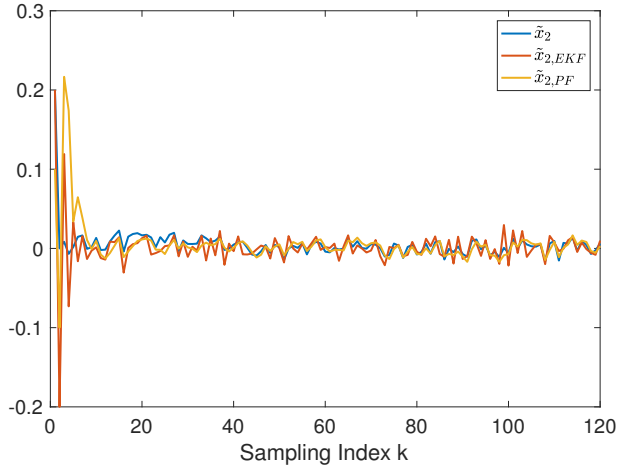
In addition,  $w$  and  $v$  are the process noise and measurement noise, respectively.

In particular,  $A_1 = A_2 = 167.4 \text{ cm}^2$  while  $c_1 = 0$  and  $c_2 = 2.88$ .  $k_0$ ,  $k_1$  and  $k_4$  are equal to 0.7, 0.25 and 0.1, respectively. The control input  $u_2$  is set as 30 while the equilibrium points are 0.23 and 0.26 for both tanks. Moreover,  $w_k$  is subjected to Gaussian distribution with zero means and 0.1 variance, meanwhile the non-Gaussian noise  $v_k$  is given obeying the following PDF:

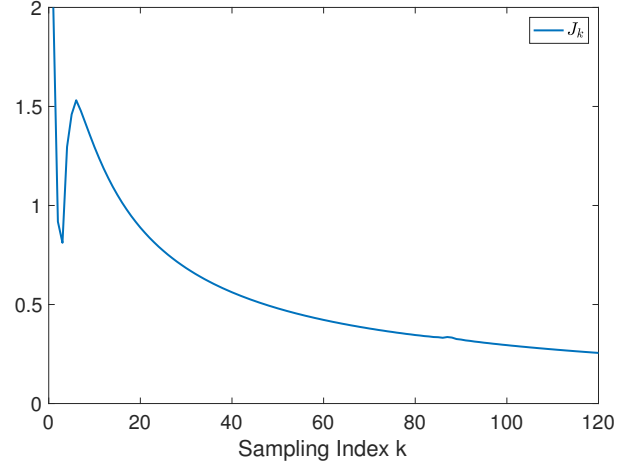
$$\gamma(\omega) = \begin{cases} 5^{12} \left( \int_0^1 \omega^8 (1-\omega)^3 d\omega \right)^{-1} \omega^8 (5-\omega)^3, & \omega \in [0, 5] \\ 0, & \text{otherwise} \end{cases} \quad (52)$$

Based on the formula, we notice that the mean-value of  $\omega_k$  is nonzero. Without loss of generality, the noise  $v_k$  in the simulation is processed as  $v_k = \omega_k - E(\omega)$  in order to shift the noise mean-value.

Following the presented filtering algorithm (3) and (31), the simulation results are given by Figs 3-6. Since the system output  $y_k$  is measurable, the entropy of  $\tilde{y}_k$  can be estimated by KDE which implies the reference signal  $r_k$ . Then the identification can be progressed to obtain the RBFNN model which leads to the filter design signal  $g_k$ . As the investigated stochastic nonlinear system has two states, the filtering results have been shown by Fig.3 and Fig.4 separately. It shows that both estimates  $\hat{x}_1$  and  $\hat{x}_2$  are very close to true values and the performances are better than standard EKF design. In addition, a further comparison for  $x_2$  is given to indicate the estimation errors in Fig.5 while the performance presented algorithm is close to particle filter design with 50 particles. Meanwhile, the performance criterion  $J$  is also given by Fig.6 while  $\lambda$  is pre-specified as 0.2. Based on the theoretical analysis and simulation results, both of the estimation errors are convergent to zero and the randomness of the filtering has been attenuated. As a summary, the presented filtering algorithm is effective with desired performances.

Fig. 3. System state  $x_1$ .Fig. 4. System state  $x_2$ .Fig. 5. Estimation error of system state  $x_2$ .

*Remark 1:* Notice that the RBFNN model training is based on output data approximation which explains the entropy

Fig. 6. Value of the performance criterion  $J$ .

increase at initial stage in Fig.6.

## VII. FURTHER DISCUSSION

This paper provides a novel design of the nonlinear filter using the RBFNN-based entropy model, and this presented model can also be applied to other research problems such as the entropy assignment and probability density function observation.

For the non-Gaussian stochastic systems, the method based on the output entropy is a significant extension comparing to the variance with the Gaussian distribution. Therefore, randomness of the systems can be described as entropy assignment problem which is similar to the co-variance assignment problem [20]. Once the nonlinear system is remodeled by the RBFNN-based entropy approximation, the entropy of the system output can be assigned as the tracking problem of the linear system. Then the performance criteria can be replaced by

$$J = \frac{1}{8} (H_y - H_r)^2 + \frac{\lambda}{2} u^2 \quad (53)$$

where  $H_r$  denotes the reference of the output entropy. Based on this criteria, the presented approach can be used to design the nonlinear control input.

Next, the output probability density function is different to estimated in real-time due to the complexity of solving the partial differential equation. In this situation, is it possible to observe the measurable output probability density function? Using the presented entropy model, the observer for output probability density function can be designed, which is similar to the observer design of the linear system.

The observer can be given as

$$\hat{V}_{k+1} = A\hat{V}_k + Bu - L \left( e^{\frac{1}{2}H_{y,k}} - e^{\frac{1}{2}\hat{H}_{y,k}} \right) \quad (54)$$

where  $L$  is the observation gain while

$$\hat{H}_{y,k} = 2 \log \left( \bar{G}\hat{V}_k \right) \quad (55)$$



To design the gain  $L$  to stabilize the observation error  $\tilde{V} = \bar{V} - \hat{V}$ , the closed-loop system is rewritten as

$$\tilde{V}_{k+1} = (A - L\bar{G})\tilde{V}_k \quad (56)$$

then the gain  $L$  can be obtained to make the matrix  $A - L\bar{G}$  Hurwitz, which means

$$\lim_{k \rightarrow \infty} \sqrt{\hat{\gamma}(y, u, L)} = \lim_{k \rightarrow \infty} C\hat{V}_k = \sqrt{\gamma(y, u)} \quad (57)$$

Therefore, the weight update equation can be considered as the online update law for the probability density function estimation which can reduce the computational complexity and guarantee the rapidity of the observation.

### VIII. CONCLUSION

In this paper, a novel nonlinear filter design has been presented via RBFNN-based entropy model. Based on this model, the dynamic relationship between the estimation error and the nonlinear filter has been established. The filtering design is transformed to the optimal input design for the transformed entropy model where the Lyapunov design method has been used and the minimum entropy of the estimation error has been achieved. Moreover, the convergence of the filtering has been analyzed while the practical system model based simulation results show the effectiveness of the presented filtering algorithm.

### ACKNOWLEDGMENT

The authors would like to thank the editor and reviewers for their valuable comments.

### REFERENCES

- [1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [2] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–108, 1961.
- [3] L. Ljung, "Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems," *IEEE Transactions on Automatic Control*, vol. 24, no. 1, pp. 36–50, 1979.
- [4] S. J. Julier and J. K. Uhlmann, "New extension of the kalman filter to nonlinear systems," in *Signal Processing, Sensor Fusion, and Target Recognition VI*, vol. 3068. International Society for Optics and Photonics, 1997, pp. 182–194.
- [5] J. L. Maryak, J. C. Spall, and G. L. Silberman, "Uncertainties for recursive estimators in nonlinear state-space models, with applications to epidemiology," *Automatica*, vol. 31, no. 12, pp. 1889–1892, 1995.
- [6] H. Wang, "Robust control of the output probability density functions for multivariable stochastic systems with guaranteed stability," *IEEE Transactions on Automatic Control*, vol. 44, no. 11, pp. 2103–2107, 1999.
- [7] —, *Bounded Dynamic Stochastic Systems: Modelling and Control*. Springer Science & Business Media, 2012.
- [8] —, "Minimum entropy control of non-gaussian dynamic stochastic systems," *IEEE Transactions on Automatic Control*, vol. 47, no. 2, pp. 398–403, 2002.
- [9] L. Guo and H. Wang, "Minimum entropy filtering for multivariate stochastic systems with non-gaussian noises," in *American Control Conference, 2005. Proceedings of the 2005*. IEEE, 2005, pp. 315–320.
- [10] L. Guo, L. Yin, H. Wang, and T. Chai, "Entropy optimization filtering for fault isolation of nonlinear non-gaussian stochastic systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 4, pp. 804–810, 2009.
- [11] J. C. Principe, *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*. Springer Science & Business Media, 2010.
- [12] P. Afshar, F. Yang, and H. Wang, "Ilc-based minimum entropy filter design and implementation for non-gaussian stochastic systems," *IEEE Transactions on Control Systems Technology*, vol. 20, no. 4, pp. 960–970, 2012.
- [13] Y. Liu, H. Wang, and C. Hou, "Ukf based nonlinear filtering using minimum entropy criterion," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 4988–4999, 2013.
- [14] J. Zhang, Y. Kuai, M. Ren, Z. Luo, and M. Lin, "Event-triggered distributed filtering for non-gaussian systems over wireless sensor networks using survival information potential criterion," *IET Control Theory & Applications*, vol. 10, no. 13, pp. 1524–1530, 2016.
- [15] B. Li, C. Liu, and W.-H. Chen, "An auxiliary particle filtering algorithm with inequality constraints," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4639–4646, 2017.
- [16] J. Zhou, H. Yue, and H. Wang, "Shaping of output pdf based on the rational square-root b-spline model," *ACTA Automatica Sinica*, vol. 31, no. 3, p. 343, 2005.
- [17] M. Brown and C. J. Harris, *Neurofuzzy adaptive modelling and control*. Prentice Hall, 1994.
- [18] T. C. Hsia, "System identification: Least-squares methods," 1977.
- [19] Y. Zhou, Q. Zhang, H. Wang, P. Zhou, and T. Chai, "EKF-based enhanced performance controller design for nonlinear stochastic systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 4, pp. 1155–1162, 2018.
- [20] Q. Zhang, Z. Wang, and H. Wang, "Parametric covariance assignment using a reduced-order closed-form covariance model," *Systems Science & Control Engineering*, vol. 4, no. 1, pp. 78–86, 2016.