Research article

Performance enhanced Kalman filter design for non-Gaussian stochastic systems with data-based minimum entropy optimisation

Qichun Zhang*

Department of Computer Science, University of Bradford, Bradford, BD7 1DP, UK

* **Correspondence:** Email: q.zhang17@bradford.ac.uk.

Abstract: Almost all of the complex dynamic processes are subjected to non-Gaussian random noises which leads to the performance deterioration of Kalman filter and Extended Kalman filter (EKF). To enhance the filtering performance, this paper presents an EKF-based filtering algorithm using minimum entropy criterion for a class of stochastic non-linear systems subjected to non-Gaussian noises. For practical implementations, the Kalman filters are widely used and the structure will not be changed due to the system integration, therefore, it is important to enhance the performance without changing the existing system design. In particular, a compensative framework has been developed where the EKF design meets the basic filtering requirements and the polynomial-based non-linear compensation has been used to adjusted the basic estimation from EKF with the entropy criterion. Since the entropy of the system output estimation error can be approximated using the measured data by kernel density estimation (KDE). A data-based framework can be obtained to enhance the performance. In addition, the presented algorithm is analysed from the view of the estimation convergence and a numerical example has been given to demonstrate the effectiveness.

Keywords: stochastic non-Gaussian systems; minimum entropy optimisation; performance enhancement; extended Kalman filter; polynomial compensation; kernel density estimation

1. Introduction

In order to control and optimise the system performance, the system state observation is a significant concept since the system state can reflect the internal properties of the dynamic system. However, for most of the practical systems, the system state cannot be measured directly, meanwhile, the random noise would affect the state and the measurement. To estimate the system state using system output measurement, Kalman filter has been presented in 1960 [3].

Kalman filter is well-developed for linear stochastic systems which are subjected to Gaussian noises. To deal with the non-linearities of the stochastic systems, extended Kalman filter (EKF) [6] has been

presented by adding the linearisation operation into the Kalman filter design. In this case, the non-linear system model has been converted to linear model locally for each instant. Due to the fact that the Kalman filter and EKF have simple parametric structure, they are widely used in practice. Notice that both Kalman filter and EKF have adopted Gaussian assumption. In other words, the performances of Kalman filter and EKF would be deteriorated if the systems are subjected to non-Gaussian noises.

Motivated by non-Gaussian distribution shaping control, the non-Gaussian systems widely exist in industrial process, for example, paper-making process, networked systems [7], neural interaction systems [16, 17], etc, it is important to investigate the filtering problem subjected to non-Gaussian noises. Note that the states or outputs will be non-Gaussian even if the noise is Gaussian due to the influences of the non-linear dynamics [8]. The performance enhancement problem for non-Gaussian control system design has been analysed in [19, 20], where the entropy criterion has been adopted to enhance the performance of the exiting fixed control loop. Similarly, this approach can be extended to filtering problem in order to overcome the fore-mentioned shortcomings.

In particular, the filter design cannot be adjusted anymore once the system integration is completed, which means the filter cannot adapt to the model time-variant dynamics, unexpected damage, non-Gaussian noise, etc. Therefore, the performance enhancement strategy should be considered based on the existing system structure and a new compensative framework for the estimated states should be established. We assume that the existing EKF design has been completed for the system and a compensative non-linear term is added to the estimated state directly. Following this approach, the residuals and un-modelled dynamics can be compensated by the non-linear term. In particular, our objective is to design an optimal non-linear term to enhance the performance.

In this paper, the non-linear term has been re-expressed by polynomial while the coefficients of the polynomial can be optimised in real time to minimise the performance criterion, where entropy has been used as performance criterion to describe the randomness of the estimation error [4]. Comparing with the existing minimum entropy filter, the convergence analysis of the presented algorithm would be obtained based upon the existing EKF structure, while the non-Gaussian effects will be attenuated by entropy optimisation. Moreover, the entropy can be estimated by kernel density estimation (KDE) with data. As a result, the analytical evolution is not essential. Different from the Kalman framework, not only the variance has been considered in the design but also the whole distribution of the estimation error has been taken into account which implies that our design includes more information and makes the performance enhancement feasible. Since the presented filtering algorithm can improve the estimation accuracy, it would affect a lot of related research topics such as fault diagnosis [5], probabilistic decoupling [15, 18] and even non-linear filtering problem [13].

The rest of this paper is organised as follows: in Section 2, the preliminaries are briefly introduced including the problem formulation, EKF and KDE. Section 3 shows the main algorithm where the compensation loop has been designed and the optimisation has been achieved. In Section 4, the convergence of the presented algorithm has been analysed. To validate the effectiveness of the presented filtering framework, the simulation results have been illustrated in Section 5. As the last part of the paper, further discussion and conclusions have been given by Section 6 and Section 7, respectively.

2. Preliminaries

Consider the stochastic non-linear system model:

$$x_{k+1} = f(x_k) + w_k$$
$$y_k = h(x_k) + v_k \tag{2.1}$$

where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$ denote the system state vector and system output, respectively. $w \in \mathbb{R}^n$ and $v \in \mathbb{R}$ stand for the random noises with arbitrary stochastic distributions where $E\{w_k\} = 0$ and $E\{v_k\} = 0$. k is the sampling index. $f : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}$ are known differentiable non-linear functions where $f(0) = 0$ and $h(0) = 0$. Note that $E\{\cdot\}$ denotes the operation for calculating the mean value.

Remark 1. *For the filtering problem, single-output means the limited measurable information which demonstrates the benefit of the system state estimation. Actually, the multi-output model can be covered using the presented algorithm with the generalised joint-entropy minimum optimisation. Without loss of the generality, the single-output model is adopted in this manuscript for simplifying the formulation.*

2.1. Extended Kalman filter

Based upon the structure of the system model, the EKF design can be started with linearisation as follows:

$$\{F_k, H_k\} = \left\{ \frac{\partial f}{\partial x}, \frac{\partial h}{\partial x} \right\} \bigg|_{x=x_k} \tag{2.2}$$

After system linearisation, the investigated system model (2.1) can be re-presented as follows:

$$x_{k+1} = F_k x_k + \bar{f}(x_k) + w_k$$
$$y_k = H_k x_k + \bar{h}(x_k) + v_w \tag{2.3}$$

where $\bar{f}(x_k)$ and $\bar{h}(x_k)$ stand for the un-modelled dynamics of the investigated system model (2.1).

The EKF filter structure can be given by Eq. (2.4)

$$\hat{x}_{k+1} = f(\hat{x}_k) + K_k \tilde{y}_k \tag{2.4}$$

where $\hat{x}_k$ denotes the estimated system state x_k for sampling index k, $\tilde{y}_k = y_k - h(\hat{x}_k)$ stand for the estimation error of the system output y_k and the filter gain $K_k \in \mathbb{R}^{n \times 1}$ can be calculated recursively as follows:

$$K_k = F_k P_k H_k^T \left(H_k P_k H_k^T + R_k \right)^{-1} \tag{2.5}$$

while the P matrix is governed by the following Riccati equation.

$$P_{k+1} = F_k P_k F_k^T + Q_k - F_k P_k H_k^T K_k^T \tag{2.6}$$

Note that R_k and Q_k stand for the covariance matrix for w_k and v_k which implies these matrices are time-varying symmetric positive definite matrices and

$$R_k = E\left\{ w_k w_k^T \right\}, Q_k = E\left\{ v_k v_k^T \right\} \tag{2.7}$$

The filtering design objective can be formulated as

$$\lim_{k \to \infty} \tilde{x}_k = 0, \forall x_0 \in \mathbb{R}^n \tag{2.8}$$

where $\tilde{x}_k = x_k - \hat{x}_k$ denote the estimation error of the system state x_k and $\hat{x}_0$ is the initial value of the filter.

For theoretical analysis, the following assumptions can be made for the investigated system model (2.1).

Assumption 1. *There exists one positive real number $L_{1,k} < 1$, such that the inequality (2.9) holds for any sampling instant k.*

$$\left\| \bar{f}(x_k) - \bar{f}(\hat{x}_k) \right\| \le L_{1,k} \|\tilde{x}_k\| \tag{2.9}$$

where $\|\cdot\|$ denotes the norm operation.

Assumption 2. *There exists one positive real number $L_{2,k} < 1$, such that the inequality (2.10) holds for any sampling instant k.*

$$\left\| \bar{h}(x_k) - \bar{h}(\hat{x}_k) \right\| \le L_{2,k} \|\tilde{x}_k\| \tag{2.10}$$

Remark 2. *Since the noises are unmeasurable, the covariance matrices R_k and Q_k cannot be obtained accurately, while the approximation also results in the performance decay. It affects the performance strongly for non-Gaussian systems because the non-Gaussian distribution cannot be characterised sufficiently by 1st-order and 2nd-order moments.*

2.2. Kernel density estimation

Note that Kalman-like filters are based on quadratic cost function which means that the optimal result is obtained using 2nd-order moment such as variance and covariance. However, since the variance and covariance cannot be used to deal with the non-Gaussian distribution properly, the concept of entropy can be introduced to describe the signal randomness which is subjected to the non-Gaussian noise. In particular, the quadratic Rényi entropy has been used in this paper.

$$H_2(\bar{\omega}) = -\log \int_\Omega \gamma^2(\bar{\omega}) d\bar{\omega} \tag{2.11}$$

where $\bar{\omega}$ denotes random variable, Ω denotes the sample space and $\gamma(\cdot)$ stands for the probability density function (PDF). Notice that the entropy is equivalent to variance for Gaussian random variable.

Once the samples of the random variable $\bar{\omega}$ can be collected as a data set $\{\bar{\omega}_1, \bar{\omega}_2, \ldots, \bar{\omega}_N\}$, then the PDF of $\bar{\omega}$ can be estimated by kernel density estimation (KDE) as follows:

$$\hat{\gamma}(\bar{\omega}) = \frac{1}{N} \sum_{i=1}^{N} G_\Sigma(\bar{\omega} - \bar{\omega}_i) \tag{2.12}$$

where G denotes the Gaussian kernel function. Σ stands for the covariance matrix of the kernel function.

Substituting the KDE equation into the definition of Rényi entropy, the estimation of entropy can be formulated as follows:

$$\hat{H}_2(\bar{\omega}) = -\log\left(\frac{1}{N^2}\sum_{i,j=1}^{N} G_{\sqrt{2}\Sigma}\left(\bar{\omega}_i - \bar{\omega}_j\right)\right) \tag{2.13}$$

Note that logarithm is a monotonic function, the entropy optimisation can be rewritten as information potential optimisation equivalently, where the information potential has been defined as

$$\hat{P}(\bar{\omega}) = \frac{1}{N^2}\sum_{i,j=1}^{N} G_{\sqrt{2}\Sigma}\left(\bar{\omega}_i - \bar{\omega}_j\right) \tag{2.14}$$

3. Performance enhanced filtering algorithm using polynomial compensation

Based upon the analysis above, the accuracy will be improved if a compensative term can be added to the EKF filter, where the EKF has been design as fixed component in terms of structure and parameters. The additional term can be inserted onto the existing filter directly which is shown as follows:

$$\hat{x}_{k+1} = f(\hat{x}_k) + K_k(y_k - h(\hat{x}_k)) + g(\tilde{y}_k) \tag{3.1}$$

Since the non-linear function $g(\tilde{y}_k)$ is difficult to estimate for sampling instant k, the polynomial form can be adopted as follows:

$$g(\tilde{y}_k) = \alpha_0 + \alpha_1\tilde{y}_k + \alpha_2\tilde{y}_k^2 + \cdots + \alpha_i\tilde{y}_k^i + \cdots \tag{3.2}$$

where α_i denotes the coefficient of the polynomial. Note that $g(0) = \alpha_0$ which results in $\alpha_0 = 0$ since $g(0) = 0$. Potentially, the polynomial can be simplified with $\alpha_0 = 0$.

Equivalently, the polynomial can be re-written based on parametric structure with optimisation purpose.

$$g(\tilde{y}_k) = \Lambda_k\phi_k(\tilde{y}_k) \tag{3.3}$$

where

$$\Lambda_k = [\alpha_0, \alpha_1, \ldots, \alpha_i, \ldots] \tag{3.4}$$

$$\phi_k(\tilde{y}_k) = \left[1, \tilde{y}_k, \tilde{y}_k^2 \ldots, \tilde{y}_k^i, \ldots\right]^T \tag{3.5}$$

while i denotes the order for polynomial estimation. Generally, high-order polynomial will lead to better estimation however the order should be pre-specified considering the computational load. Without loss of generality, we can choose m as the order of the estimation. Notice that $m = 1$ results in a linear compensator. In addition, Eq. (3.3) can be further simplified since $g(0) = 0$. As a result, $\alpha_0 = 0$ while the first element can be removed from Eq. (3.5) and α_0 can be also removed from Eq. (3.4).

Since ϕ_k is known for any sampling instant k, the parameter matrix Λ_k should be obtained with the entropy optimisation. Using KDE and information potential definition, we use information potential to describe the randomness while it can be approximated as follows:

$$\hat{P}_2 = \left(\frac{1}{k^2} \sum_{i,j=1}^{k} G_{\sqrt{2}\Sigma} \left(\tilde{y}_i - \tilde{y}_j \right) \right) \tag{3.6}$$

where G denotes the Gaussian kernel function. Σ stands for the covariance matrix of the kernel function.

Based upon the obtained entropy value, Eq. (3.7) can be considered as the performance criterion for the optimisation.

$$J_k = R_1 \hat{P}_2^2 + R_2 \Lambda_k^T \Lambda_k + R_3 E \left\{ \tilde{y}_k^2 \right\} \tag{3.7}$$

where R_1, R_2 and R_3 are the weighting constants.

To minimise the performance criterion (3.7), the gradient descent optimisation can be adopted as follows:

$$\Lambda_{k+1} = \Lambda_k - \varepsilon_k \frac{\partial J_k}{\partial \Lambda_k} \tag{3.8}$$

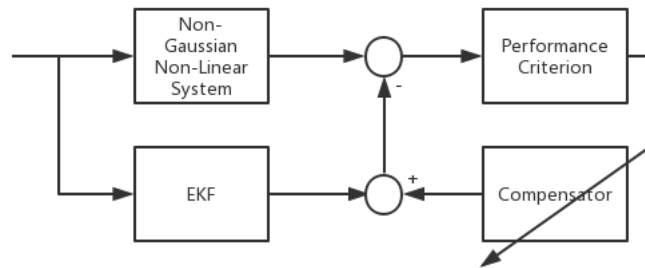where ε_k is the time-varying step for optimisation.

Figure 1. The block diagram of the presented filtering algorithm. Note that the existing EKF loop will not be affected by the additional compensative loop.

Note that there always exists a real sufficient large-value constant R_2, such that the inequality (3.9) can be satisfied.

$$\frac{\partial^2 J}{\partial \Lambda_k^2} > 0 \tag{3.9}$$

which results in the existence of the optimum.

To illustrate the presented algorithm, the block diagram is given by Figure 1.

4. Convergence

The convergence analysis of the presented filtering algorithm should be started from the dynamics of the system state estimation error. Particularly, substituting the designed filter (3.1) (3.3) into the system model (2.1), we have

$$\tilde{x}_{k+1} = f(x_k) - f(\hat{x}_k) - K_k \tilde{y}_k - \Lambda_k \phi_k(\tilde{y}_k) + w_k$$
$$= F_k \tilde{x}_k + \bar{f}(x_k) - \bar{f}(\hat{x}_k) - K_k \tilde{y}_k - \Lambda_k \phi_k(\tilde{y}_k) + w_k \tag{4.1}$$

Moreover, we have

$$\tilde{y}_k = H_k \tilde{x}_k + \bar{h}(x_k) - \bar{h}(\hat{x}_k) + v_w \tag{4.2}$$

Substituting $\tilde{y}_k$ and Λ_k into the equations (4.2) and (4.3), we have

$$\tilde{x}_{k+1} = F_k \tilde{x}_k + \bar{f}(x_k) - \bar{f}(\hat{x}_k) + w_k$$
$$- K_k \left(H_k \tilde{x}_k + \bar{h}(x_k) - \bar{h}(\hat{x}_k) + v_w \right) - \left(\Lambda_{k-1} - \varepsilon_{k-1} \frac{\partial J_{k-1}}{\partial \Lambda_{k-1}} \right) \phi_k(\tilde{y}_k)$$
$$= (F_k - K_k H_k) \tilde{x}_k + \bar{f}(x_k) - \bar{f}(\hat{x}_k)$$
$$- K_k \left(\bar{h}(x_k) - \bar{h}(\hat{x}_k) \right) - K_k v_k + w_k + \left(\Lambda_{k-1} - \varepsilon_{k-1} \frac{\partial J_{k-1}}{\partial \Lambda_{k-1}} \right) \phi_k(\tilde{y}_k) \tag{4.3}$$

Using Assumption A2, the inequality (4.4) can be obtained.

$$\|\tilde{y}_k\| = \|H_k\| \|\tilde{x}_k\| + \left\| \bar{h}(x_k) - \bar{h}(\hat{x}_k) \right\| + \|v_k\|$$
$$\leq \|H_k\| \|\tilde{x}_k\| + L_{2,k} \|\tilde{x}_k\| + \|v_k\| \tag{4.4}$$

which leads to

$$E\{\|\tilde{y}_k\|\} \leq (\|H_k\| + L_{2,k}) E\{\|\tilde{x}_k\|\} \tag{4.5}$$

Using the structure of ϕ_k, there always exist two real positive numbers $\bar{L}_{2,k}$ and $\bar{\sigma}_k$, such that,

$$E\{\|\phi_k(\tilde{y}_k)\|\} \leq \bar{L}_{2,k} E\{\|\tilde{x}_k\|\} + \bar{\sigma}_k \tag{4.6}$$

where $\bar{\sigma}_k$ denotes the upper bound for α_0 for all sampling instants.

Substituting Eqs. (4.5) and (4.6) to Eq. (4.3) with Assumption A1, we have

$$E\{\|\tilde{x}_{k+1}\|\} = (F_k - K_k H_k) E\{\|\tilde{x}_k\|\} + E\{w_k\} + \left\| \bar{f}(x_k) - \bar{f}(\hat{x}_k) \right\| + E\{K_k v_k\}$$
$$+ \|K_k\| E\left\{ \left\| \bar{h}(x_k) - \bar{h}(\hat{x}_k) \right\| \right\} + \left\| \Lambda_{k-1} - \varepsilon_{k-1} \frac{\partial J_{k-1}}{\partial \Lambda_{k-1}} \right\| E\{\|\phi_k(\tilde{y}_k)\|\}$$
$$\leq \|F_k - K_k H_k\| E\{\|\tilde{x}_k\|\} + \|K_k\| L_{2,k} E\{\|\tilde{x}_k\|\} + L_{1,k} E\{\|\tilde{x}_k\|\}$$
$$+ \left\| \Lambda_{k-1} - \varepsilon_{k-1} \frac{\partial J_{k-1}}{\partial \Lambda_{k-1}} \right\| \left(\bar{L}_{2,k} E\{\|\tilde{x}_k\|\} + \bar{\sigma}_k \right)$$
$$= \left(\|F_k - K_k H_k\| + L_{1,k} + L_{2,k} \|K_k\| + \bar{L}_{2,k} \left\| \Lambda_{k-1} - \varepsilon_{k-1} \frac{\partial J_{k-1}}{\partial \Lambda_{k-1}} \right\| \right) E\{\|\tilde{x}_k\|\}$$
$$+ \bar{L}_{2,k} \left\| \Lambda_{k-1} - \varepsilon_{k-1} \frac{\partial J_{k-1}}{\partial \Lambda_{k-1}} \right\| \bar{\sigma}_k \tag{4.7}$$

Note that the estimation error $\tilde{x}_k$ is bounded if

$$\|F_k - K_k H_k\| + L_{1,k} + L_{2,k} \|K_k\| + \bar{L}_{2,k} \left\|\Lambda_{k-1} - \varepsilon_{k-1}\frac{\partial J_{k-1}}{\partial \Lambda_{k-1}}\right\| < 1 \tag{4.8}$$

In addition, the selected step ε_{k-1} makes the inequality (4.9) holds, naturally.

$$\bar{L}_{2,k} \left\|\Lambda_{k-1} - \varepsilon_{k-1}\frac{\partial J_{k-1}}{\partial \Lambda_{k-1}}\right\| < 1 \tag{4.9}$$

which leads that the upper bound $\bar{\sigma}_k$ is decreasing along k. In other words, $\bar{\sigma}_k \to 0$ implies that the estimation error converges to 0 in mean-norm sense. In this paper, $\bar{\sigma}_k = 0$ since $\alpha_0 = 0$ which simplifies Eq. (4.8) as follows:

$$E\{\|\tilde{x}_{k+1}\|\} = \left(\|F_k - K_k H_k\| + L_{1,k} + L_{2,k}\|K_k\| + \bar{L}_{2,k}\left\|\Lambda_{k-1} - \varepsilon_{k-1}\frac{\partial J_{k-1}}{\partial \Lambda_{k-1}}\right\|\right) E\{\|\tilde{x}_k\|\} \tag{4.10}$$

Note that $\|F_k - K_k H_k\| < 1, L_{1,k} < 1$ and $L_{2,k} < 1$, Eq. (4.8) is implementable. For any sampling instant k, the system evaluation is governed by the pre-specified step ε_{k-1}. Note that the convergence can be achieved with the information from the previous sampling instant then J_k is always bounded.

Based upon the convergence analysis, the design procedure can be demonstrated using the following pseudo-code: Algorithm 1.

Algorithm 1 Entropy-based EKF with non-linear compensation

Require: Data collection of system output y
Input: The existing EKF
Output: The measured system output y_k

   Initialization: Pre-specified the kernel density function, the dimension m for polynomial order, the time t_c to add the compensative signal into the system and the t_s for operation time.
   for $k \leq t_s$ **do**
      Collect the time sampling data set of y
      if $k \geq t_c$ **then**
         Estimate the entropy using the collected data y
         Calculate the value of cost function
         Select the step to meet the system convergence condition
         Calculate the compensative gain Λ_k using gradient descent optimisation
         K_k is fixed by EKF and add the compensator into system
      else
         K_k is fixed by EKF without compensative loop
      end if
      Update the system with system model
      $k = k + 1$
   end for

---

5. Simulation

In order to demonstrate the effectiveness of the presented algorithm especially for practical application, the two-tank system model, shown by Figure 2, has been investigated in terms of level control. In particular, the water in the pool is injected into tank 1 with controlling the pump $u_{2,k}$ and flow from tank 2 to tank 1 goes through the intercommunicating valve, while the valve opening ratio $u_{1,k}$ and the discharge vale of tank 2 are closed. Based on [13], the system model can be formulated as follows:

$$A_1 \frac{dx_1}{dt} = k_3 u_1 - c_1 - k_1 \sqrt{x_1} + k_0 \sqrt{x_2 - x_1} + w_1$$

$$A_2 \frac{dx_2}{dt} = k_4 u_2 - c_2 - k_3 u_1 + c_1 - k_2 \sqrt{x_2} - k_0 \sqrt{x_2 - x_1} + w_2$$

$$y = x_1 + v \tag{5.1}$$

where x_1 is the level of tank 1and x_2 denotes the level of tank 2. A_1 and A_2 are the cross-sectional area. c_1 and c_2 are constant parameters of the valves and pumps. k_0, k_1, k_2, k_3 and k_4 stand for the ratio of the valves. Note that w and v are the process noise and measurement noise, respectively. Furthermore, w and v are non-Gaussian noises although they are with zero mean values.
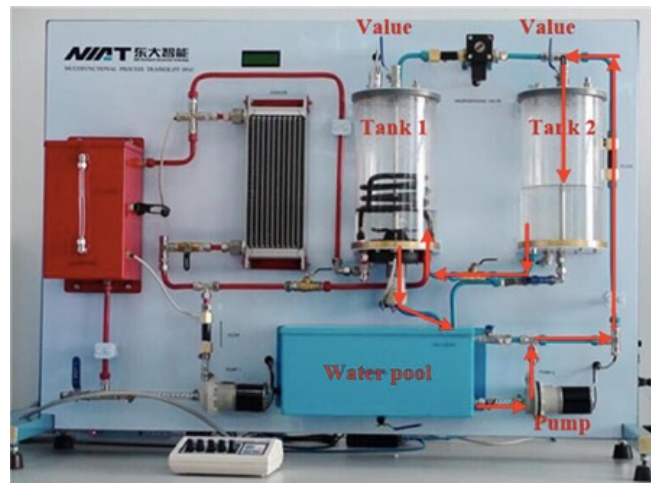

Figure 2. The investigated two-tank water level control system.

To apply the presented filtering algorithm, the discrete-time model can be obtained following the discretisation of system model (5.1) where the sampling time is selected as 0.1s. In addition, the parameters of the model can be pre-measured as $A_1 = A_2 = 167.4 cm^2$, $k_0 = 0.7$, $k_1 = 0.25$, $k_2 = k_3 = 0$, $k_4 = 0.1$, $c_1 = 0$ and $c_2 = 2.88$. Since the value of tank 1 is assumed as closed, we have $u_1 = 0$ while the system is driven by u_2 only where it has been pre-specified as $u_2 = 30$.

Without loss of generality, we choose $m = 2$ which results in a 2-nd order polynomial based compensator and $\alpha_0 = 0$. The weighting constants are given as $R_1 = R_3 = 1$ and $R_2 = 50$. Before $k = 3000$, only EKF is used where the EKF can give a good tracking property for any initial value, however the entropy of each instant has not been optimised. The enhanced performance scheme will be adopted based on existing EKF design after $k = 3000$ while the entropy will be attenuated by the additional compensation.

In particular, Figure 3 shows the performance of the estimation for system output, where the measurement noise v has been attenuated. And based on the estimation error $\tilde{y}_k$, the system state x_1 and x_2 can be estimated while performances have been illustrated by Figures 4 and 5. It has been shown that the non-Gaussian noise strongly affects the process. The presented algorithm can supply the good performance for system state estimation subject to the non-Gaussian noise. In addition, the estimation errors $\tilde{x}_1$ and $\tilde{x}_2$ are given by Figures 6 and 7, where the estimation errors are bounded and close to zero.
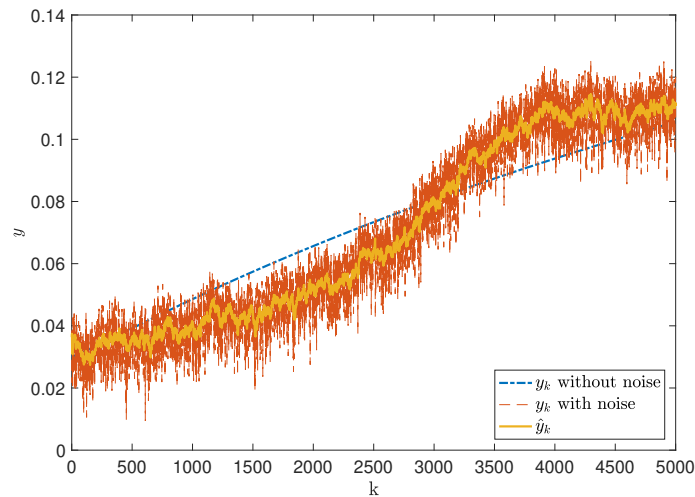
Figure 3. The measured system output y_k, the ideal dynamics of y_k without noise and the estimated system state $\hat{y}_k$.
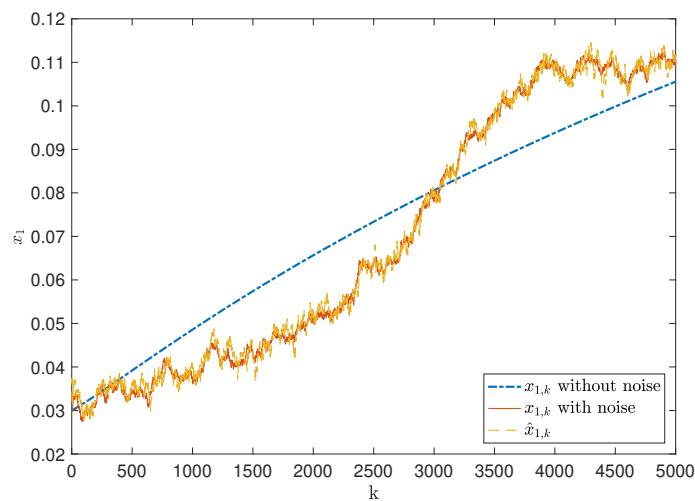
Figure 4. The measured system state $x_{1,k}$, the ideal dynamics of $x_{1,k}$ and the estimated system state $\hat{x}_{1,k}$.
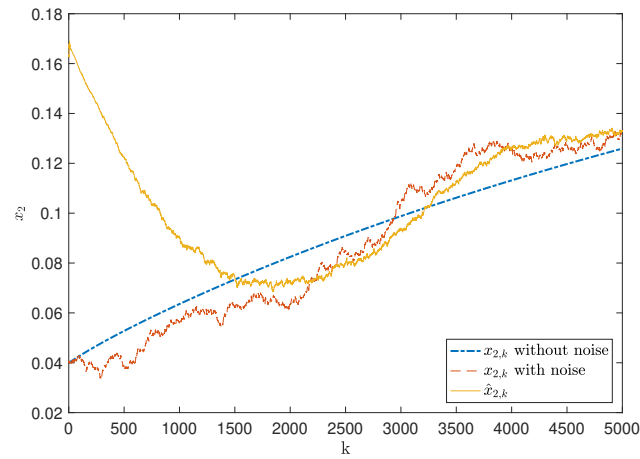
Figure 5. The measured system state $x_{2,k}$, the ideal dynamics of $x_{2,k}$ and the estimated system state $\hat{x}_{2,k}$.
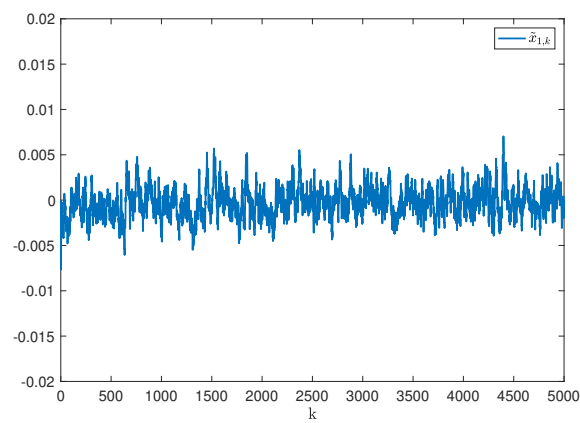

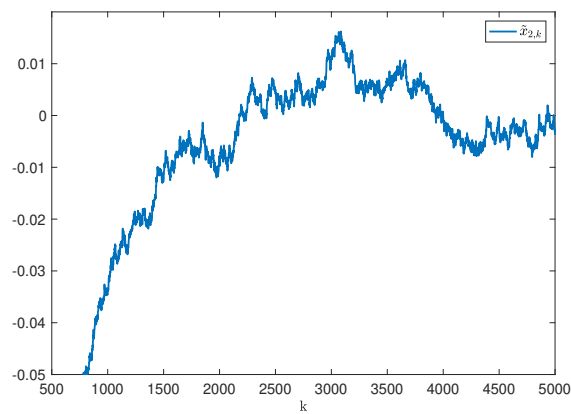**Figure 6.** The system state x_1 estimation error $\tilde{x}_1$.

Figure 7. The system state x_2 estimation error $\tilde{x}_2$.

Based upon the performance enhancement via entropy attenuation, the probability density function of the system output estimation error should become sharp which has been shown by Figure 8. Note that after $k = 3000$, the probability density function becomes sharper which implies that the entropy value has been attenuated and the randomness of the estimation error has been reduced. Meanwhile, the performance criterion is given by Figure 9 where the value of the entropy-based cost function has been minimised. Figure 8 also illustrates the comparison of the performance between the EKF and the presented filtering scheme, where the performance has been enhanced in terms of minimum randomness of the estimation. Note that the EKF can be used firstly and then the filter design can be switched to the enhancement scheme as the presented algorithm will not affect the existing design loop. Meanwhile, the convergence is much easier to achieved once the estimation error has been governed within a neighbourhood of zero point.
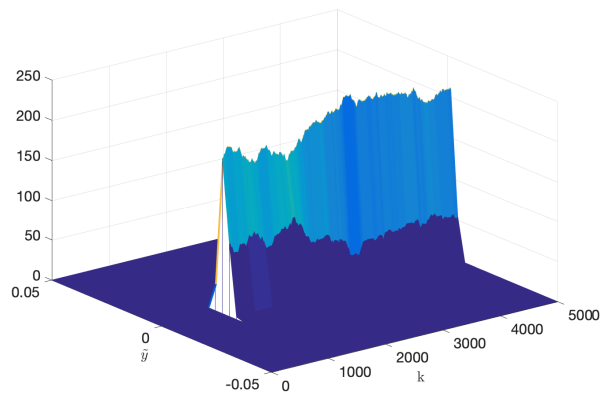

Figure 8. The probability density function of the system output estimation error $\tilde{y}_k$. Note that the sharp probability density function leads to the attenuated entropy.
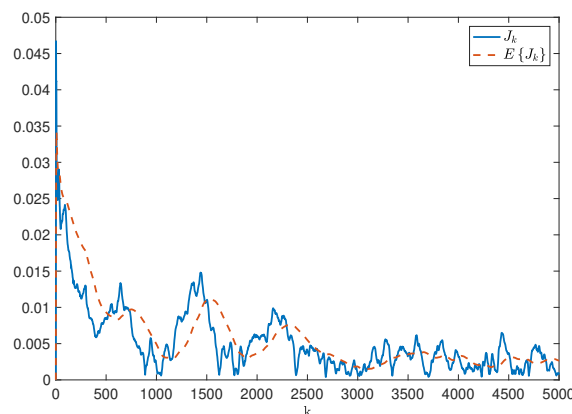

Figure 9. The value of performance criterion J_k and its mean value.

6. Further discussion

The contribution of the presented algorithm can be considered as the framework. In particular, most of the practical system structures cannot be changed after the design progress. In other words, it is significant to enhance the performance without changing the existing design from the view of the implementation. Moreover, this paper shows the framework with EKF as a standard example. The approach can be adopted with all the other filter design such as high-gain [2], NN-based filter [13], even particle filter [14], etc. On the other hand, the performance criterion can also be replaced by other optimisation objectives such as the economic dispatch [9], etc.

Notice that the presented algorithm will not increase the complexity of the algorithm implementation. For each instant, the optimisation operations have been merged into the EKF algorithm, however, only one iteration has been adopted for each instant optimisation which means the addition complexity for each instant k is equal to $O(M)$, where M is a positive integer constant denoting the operation number for optimisation at each instant. In other words, along the system runs $k \to \infty$, the additional compensator complexity will be kept at the same level comparing to the existing filter design algorithm.

Since a number of the exciting fault diagnosis approaches are based on the filter or observer designs. The presented framework has a direct perspective for practical applications. Due to the fact that the probability density function based design has been used for fault diagnosis and tolerance control [1, 12], the entropy performance enhancement can be merged into the existing methods directly. Another perspective is the potential applications for operation control of industrial processes, such as magnesium furnaces [10, 11], etc.

7. Conclusions

In this paper, a novel filtering algorithm has been proposed in order to enhance the estimation accuracy for the non-Gaussian system based upon the existing EKF design. Basically, the existing EKF design has not been changed in terms of the structure and the parameters. An additional compensation loop has been designed using the non-linear compensator in the polynomial format with adjustable coefficients. Moreover, this addition non-linear term can be restated as a gain vector with a measurable vector, then the optimal coefficients should be obtained as the gain of the compensator. To achieve this optimisation, the concept of the entropy and kernel density estimation (KDE) have been adopted where the parameters have been optimised with the minimum entropy cost function using the gradient descent approach. The theoretical analysis demonstrates that the additional compensative loop design is able to guarantee the convergence of the estimation error and the simulation results indicate the effectiveness of the presented algorithm.

Acknowledgments

Conflict of interest

The author declares no conflicts of interest in this paper.

References

1. Dai X and Gao Z (2013) From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis. *IEEE T Ind Inform* 9: 2226–2238.

2. Dai X, Gao Z, Breikin T, et al. (2012) High-gain observer-based estimation of parameter variations with delay alignment. *IEEE T Automat Contr* 57: 726–732.

3. Grewal MS (2011) Kalman filtering, Springer.

4. Guo L and Wang H (2006) Minimum entropy filtering for multivariate stochastic systems with non-gaussian noises. *IEEE T Automat Contr* 51: 695–700.

5. Guo L, Yin L, Wang H, et al. (2009) Entropy optimization filtering for fault isolation of nonlinear non-gaussian stochastic systems. *IEEE T Automat Contr* 54: 804–810.

6. Reif K, Gunther S, Yaz E, et al. (1999) Stochastic stability of the discrete-time extended kalman filter. *IEEE T Automat Contr* 44: 714–728.

7. Ren M, Zhang J, Jiang M, et al. (2015) Minimum ($\{h, \phi\}$)- entropy control for non-gaussian stochastic networked control systems and its application to a networked dc motor control system. *IEEE Transactions on Control Systems and Technology* 23: 406–411.

8. Ren M, Zhang Q and Zhang J (2019) An introductory survey of probability density function control. *Systems Science & Control Engineering* 7: 158–170.

9. Wang S, Wang H, Fan R, et al. (2018) Objective pdf-shaping-based economic dispatch for power systems with intermittent generation sources via simultaneous mean and variance minimization. *2018 IEEE 14th International Conference on Control and Automation (ICCA)*, 927–934.

10. Yang J and Chai T (2016) Data-driven demand forecasting method for fused magnesium furnaces. *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, 2015–2022.

11. Yang J, Chai T, Luo C, et al. (2019) Intelligent demand forecasting of smelting process using data-driven and mechanism model. *IEEE T Ind Electron* 66: 9745–9755.

12. Yao L, Qin J, Wang H, et al. (2012) Design of new fault diagnosis and fault tolerant control scheme for non-gaussian singular stochastic distribution systems. *Automatica* 48: 2305–2313.

13. Yin X, Zhang Q, Wang H, et al. (2019) Rbfnn-based minimum entropy filtering for a class of stochastic nonlinear systems. *IEEE T Automat Contr*, 1–1.

14. Yu M, Liu C, Li B, et al. (2016) An enhanced particle filtering method for gmti radar tracking. *IEEE T Aero Elec Sys* 52: 1408–1420.

15. Zhang Q and Hu L (2018) Probabilistic decoupling control for stochastic non-linear systems using ekf-based dynamic set-point adjustment. *2018 UKACC 12th International Conference on Control (CONTROL)*, 330–335.

16. Zhang Q and Sepulveda F (2017) Entropy-based axon-to-axon mutual interaction characterization via iterative learning identification. *EMBEC & NBC 2017*, 691–694, Springer.

17. Zhang Q and Sepulveda F (2017) A model study of the neural interaction via mutual coupling factor identification. *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 3329–3332.

18. Zhang Q and Yin X (2018) Observer-based parametric decoupling controller design for a class of multi-variable non-linear uncertain systems. *Systems Science & Control Engineering* 6: 258–267.

19. Zhou Y, Zhang Q and Wang H (2016) Enhanced performance controller design for stochastic systems by adding extra state estimation onto the existing closed loop control. *2016 UKACC 11th International Conference on Control (CONTROL)*, 1–6.

20. Zhou Y, Zhang Q, Wang H, et al. (2018) Ekf-based enhanced performance controller design for nonlinear stochastic systems. *IEEE T Automat Contr* 63: 1155–1162.