SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •
Special Focus on Analysis and Synthesis for Stochastic Systems

December 2017, Vol. 60 120203:1–120203:11
doi: 10.1007/s11432-017-9138-6

Entropy optimization based filtering for non-Gaussian stochastic systems

Bo TIAN[1], Yan WANG[1*] & Lei GUO[1,2]

[1]*School of Automation Science and Electrical Engineering, Beihang University, Beijing* 100191*, China;*
[2]*National Key Laboratory on Aircraft Control Technology, Beihang University, Beijing* 100191*, China*

Received April 18, 2017; accepted June 13, 2017; published online November 9, 2017

Abstract This paper is concerned with the entropy optimization based filter design for a class of multivariate dynamic stochastic systems with simultaneous presence of non-Gaussian process noise and measurement noise. The filter consists of time update and measurement update two steps, where the selection of the filter gain in the measurement update equation is a key issue to be addressed. Different from the classic Kalman filter theory, entropy rather than variance is employed as the filtering performance criterion due to the non-Gaussian characteristic of the estimation error. Following the establishment of the relationship between the probability density functions of random noises and estimation error, two kinds of entropy based performance indices are provided. On this basis, the corresponding optimal filter gains are obtained respectively by using the gradient optimization technique. Finally, some numerical simulations are provided to demonstrate the effectiveness of the proposed filtering algorithms.

Keywords non-Gaussian systems, joint probability density function (JPDF), quadratic information potential, relative entropy, optimal filtering

Citation Tian B, Wang Y, Guo L. Entropy optimization based filtering for non-Gaussian stochastic systems. Sci China Inf Sci, 2017, 60(12): 120203, doi: 10.1007/s11432-017-9138-6

1 Introduction

State estimation has always been an important research aspect in control, signal processing and communication communities, and has received considerable attention following the development of the Kalman filtering theory [1–5]. Under the Gaussian white noise assumption, Kalman filter provides the minimum variance estimate where the variance of estimation error indicates the filtering performance. However, for stochastic systems subjected to non-Gaussian noises, Kalman filtering algorithm may not perform well. Although several techniques such as H_∞ filtering [6, 7] and quadratic error-constrained filtering [8, 9] have been proposed and proven to be less sensitive to non-Gaussian noise, they may lead to certain restrictions and conservations because of ignoring the knowledge of noise statistics. As we know, the stochastic property of a non-Gaussian variable (especially the one with nonsymmetric distribution) cannot be purely described by the variance [10]. Therefore, some other performance criteria need to be taken into account so as to design an appropriate filter. It should be noticed that entropy has been adopted in some literatures [11, 12].

* Corresponding author (email: w-yan@buaa.edu.cn)

Entropy, as a unified probabilistic measure of uncertainty quantification, has been widely used in information theory. And it has also been introduced into the modeling, control and estimation fields [13–18]. It should be pointed out that some efforts have been made to employ the entropy as the performance index in the filter design for general non-Gaussian stochastic systems. The corresponding minimum entropy filtering algorithm was firstly proposed in [11]. After establishing several new concepts including hybrid random vector, hybrid probability and hybrid entropy, the probability density function (PDF) of estimation error was formulated in terms of the noise PDF and the filter gain matrix. Using the entropy based performance index, a recursive optimal filtering algorithm was put forward such that the hybrid entropy of the estimation error was minimized and the local stability of the error dynamics was guaranteed as well. Besides, a relative entropy based optimal tracking filter was developed for non-Gaussian stochastic systems in [12]. The key idea was to ensure the characteristic function of estimation error to follow closely a target characteristic function. For this purpose, the relationship between the hybrid characteristic functions of stochastic input and estimation error was set up. Based on the minimum relative entropy criterion, an optimal filtering algorithm with a compact form was then provided. However, it is worth mentioning that there are still some problems remained to be solved in the above two papers. Firstly, the filters proposed in [11, 12] have no explicit steps of time update and measurement update. It lacks of the detailed analysis of the statistical properties of priori and posteriori estimates. Secondly, little attention is paid to the system with simultaneous presence of process noise and measurement noise. Only the process noise is considered in [11], and the study in [12] does not make full use of the statistical property of the measurement noise.

Motivated by the above observations, this paper is focused on the filtering algorithm design for multivariate stochastic systems with simultaneous presence of non-Gaussian process noise and measurement noise by utilizing the entropy based performance criteria. The filter is formed by time update and measurement update these two steps, where the filter gain in the measurement update equation decides the filtering performance and is a key issue to be designed. In order to measure the filtering performance, the extended concepts of entropy including quadratic information potential and relative entropy are employed respectively in this paper. The error JPDFs (joint probability density functions) of priori and posteriori estimates are represented in terms of the JPDFs of noises and the undetermined filter gain, based on which the two performance indices are constructed. Following the calculation of optimal filter gain, two kinds of entropy optimization based filtering algorithms are provided.

The rest of this paper is organized as follows. The problem formulation and some preliminaries are introduced in Section 2. Section 3 presents the design procedures of the entropy based filtering algorithms. Numerical examples together with the simulation results are included in Section 4. Finally, we conclude the paper in Section 5.

The notations used in this paper are standard. In the following, if not stated, matrices are assumed to have appropriate dimensions. The identity matrix and zero matrix are denoted by I and 0, respectively. For a symmetric matrix P, the notation $P \geqslant 0$ means that P is positive semidefinite. ∇ denotes the gradient of a function. $\mathrm{E}\{\cdot\}$ stands for the mathematical expectation of a random variable (vector), and $\gamma(\cdot)$ denotes the (joint) probability density function.

2 Problem formulation and preliminaries

Consider the following linear stochastic time-varying system:

$$\begin{cases} x_k = A_{k-1}x_{k-1} + \omega_{k-1}, \\ z_k = C_k x_k + \nu_k, \end{cases} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the system state, $z_k \in \mathbb{R}^m$ is the measurement output, $\omega_k \in \mathbb{R}^n$ and $\nu_k \in \mathbb{R}^m$ are the process noise and measurement noise respectively. A_k and C_k are two known time-varying system matrices. Different from the classic Kalman filtering theory, the noises in the system are arbitrary random vectors not confined to the Gaussian case. In fact, non-Gaussian noises exist widely in many practical systems. A typical example is the integrated navigation system, where the involved noises are often of non-Gaussian types. Still, it is encouraging that with the development of instrumentation, computer science and data processing technology, some identification methods have been developed to model the PDFs of random variables, such as the kernel estimation technique, direct physical measurement, and other experimental techniques [19, 20]. Thus, the following assumptions are made in this paper, which can be met in many practical systems.

Assumption 1. The process noise ω_k and the measurement noise ν_k are mutually independent, and each follows a known probability distribution with zero mean. Their JPDFs are denoted by $\gamma_{\omega_k}(\tau)$ and $\gamma_{\nu_k}(\tau)$, which are defined on $[\alpha_k, \beta_k]^n$ and $[\delta_k, \varepsilon_k]^m$ respectively.

Assumption 2. The initial state x_0 is independent of ω_k and ν_k, and is with a known JPDF denoted by $\gamma_{x_0}(\tau)$.

Remark 1. The system considered in this paper is subjected to non-Gaussian process noise and non-Gaussian measurement noise simultaneously. For the system discussed in [11], only process noise is considered. And the work in [12] only makes use of the statistical property (characteristic function) of process noise, and no information except the known bounded mean of measurement noise is utilized for the filter design. In this paper, we investigate the system with simultaneous presence of non-Gaussian process noise and measurement noise, and such a presence is quite typical in engineering.

The purpose of filter design is to use available information of the system to estimate the state x_k. For the dynamic system given by (1), the filter is proposed with a classic recursive form:

- Time update (predict)

$$\hat{x}_k^- = A_{k-1}\hat{x}_{k-1}^+, \tag{2}$$

- Measurement update (correct)

$$\hat{x}_k^+ = \hat{x}_k^- + L_k(z_k - C_k\hat{x}_k^-), \tag{3}$$

where $\hat{x}_k^-$ is the priori estimate of x_k, $\hat{x}_k^+$ is the posteriori estimate of x_k, and $L_k \in \mathbb{R}^{n\times m}$ is the filter gain matrix to be determined. Combining (2) and (3) into a single equation yields the one-step filter with the following form:

$$\begin{aligned}\hat{x}_k^+ &= A_{k-1}\hat{x}_{k-1}^+ + L_k(z_k - C_kA_{k-1}\hat{x}_{k-1}^+)\\ &= (I - L_kC_k)A_{k-1}\hat{x}_{k-1}^+ + L_kz_k.\end{aligned} \tag{4}$$

And the estimation of the initial state is taken to be $\hat{x}_0^+ = \mathrm{E}\{x_0\}$.

Define the estimation error as $e_k = x_k - \hat{x}_k^+$. A desired filter should ensure that a measure of e_k is optimized. Due to the non-Gaussian characteristics of input noises, the estimation error e_k is also a non-Gaussian random variable, and the variance is not sufficient to characterize its stochastic property. In this case, it is nature to consider some other criteria rather than the variance so as to produce the satisfactory filtering effect. Generally speaking, the filter design objective should be to minimize the randomness of the estimation error. It has been well known that entropy is a unified probabilistic measure of uncertainty quantification. Therefore, entropy and its extended concepts (such as information potential, relative entropy and so on) can be adopted to form the criteria for the filter design for non-Gaussian systems. The Shannon entropy of e_k is defined as follows [21]:

$$H(e_k) = -\int_{[a_k,b_k]^n} \gamma_{e_k}(\tau)\ln\gamma_{e_k}(\tau)\mathrm{d}\tau, \tag{5}$$

where $\gamma_{e_k}(\tau)$ is the JPDF of the estimation error defined on $[a_k, b_k]^n$. Other types of entropy and several extended concepts will be introduced in the following discussion.

The filter design is performed by optimizing the entropy, which means that the shape of the error PDF is made as narrow as possible and the randomness of estimation error is minimized. Therefore, the filter proposed in this paper is called entropy optimization based filter. The specific design procedures and performance analysis will be presented in the next section.

3 Main results

3.1 JPDF formulation of estimation error

To start with, we focus on the derivation of the estimation error dynamics. Denote the prediction error as s_k. From (1) and (2), we can get

$$s_k = x_k - \hat{x}_k^- = A_{k-1}e_{k-1} + \omega_{k-1}. \tag{6}$$

Furthermore, we have

$$\begin{aligned} e_k &= x_k - \hat{x}_k^+ \\ &= x_k - \hat{x}_k^- - L_k(C_k x_k + \nu_k - C_k\hat{x}_k^-) \\ &= s_k - L_k(C_k s_k + \nu_k) \\ &= (I - L_k C_k)s_k - L_k\nu_k. \end{aligned} \tag{7}$$

Substituting (6) into (7) yields the dynamics of the estimation error as follows:

$$e_k = (I - L_kC_k)A_{k-1}e_{k-1} + (I - L_kC_k)\omega_{k-1} - L_k\nu_k. \tag{8}$$

As $e_0 = x_0 - \hat{x}_0^+ = x_0 - \mathrm{E}\{x_0\}$, it is easy to observe that $\mathrm{E}\{e_0\} = 0$ and $\gamma_{e_0}(\tau) = \gamma_{x_0}(\tau + \mathrm{E}\{x_0\})$. According to Assumptions 1 and 2, we can conclude $\mathrm{E}\{e_k\} = 0$ holds, which means that the estimation is unbiased.

In order to calculate the entropy, the JPDF of the estimation error should be formulated in advance, which is presented in the following theorem.

Theorem 1. Under Assumptions 1 and 2, at sample time k, the JPDF of s_k is given by

$$\gamma_{s_k}(\eta) = \int_{[a_{k-1},b_{k-1}]^n} \gamma_{e_{k-1}}(\rho)\gamma_{\omega_{k-1}}(\eta - A_{k-1}\rho)\mathrm{d}\rho. \tag{9}$$

And furthermore, the JPDF of e_k can be formulated by

$$\gamma_{e_k}(\tau) = \int_{[\delta_k,\varepsilon_k]^m} \gamma_{s_k}\left[(I - L_kC_k)^{-1}(\tau + L_k\sigma)\right]\gamma_{\nu_k}(\sigma)\left|\det(I - L_kC_k)^{-1}\right|\mathrm{d}\sigma. \tag{10}$$

Proof. From (6), we see that the JPDF of s_k can be formulated in terms of the JPDFs of e_{k-1} and ω_{k-1}. For this purpose, an auxiliary vector is defined as

$$\bar{s}_k = \begin{bmatrix} s_k \\ e_{k-1} \end{bmatrix} = \begin{bmatrix} A_{k-1}e_{k-1} + \omega_{k-1} \\ e_{k-1} \end{bmatrix} = \begin{bmatrix} I & A_{k-1} \\ 0 & I \end{bmatrix}\begin{bmatrix} \omega_{k-1} \\ e_{k-1} \end{bmatrix}. \tag{11}$$

It can be seen that

$$\gamma_{s_k}(\eta) = \int_{[a_{k-1},b_{k-1}]^n} \gamma_{\bar{s}_k}(\eta,\rho)\mathrm{d}\rho, \tag{12}$$

where $[a_{k-1}, b_{k-1}]^n$ is the interval of e_{k-1}. For the defined auxiliary mapping (11), applying the fundamental PDF transformation property leads to

$$\begin{aligned} \gamma_{\bar{s}_k}(\eta,\rho) &= \gamma_{\omega_{k-1},e_{k-1}}(\eta - A_{k-1}\rho, \rho)\left|\det\begin{bmatrix} I & A_{k-1} \\ 0 & I \end{bmatrix}\right|^{-1} \\ &= \gamma_{\omega_{k-1}}(\eta - A_{k-1}\rho)\gamma_{e_{k-1}}(\rho), \end{aligned} \tag{13}$$

where the second equality is obtained due to the independence of ω_{k-1} and e_{k-1}. Substituting (13) into (12) yields the JPDF of s_k as shown in (9).

Next, we can formulate the JPDF of e_k in terms of those of s_k and ν_k based on (7). Construct the following auxiliary vector

$$\bar{e}_k = \begin{bmatrix} e_k \\ \nu_k \end{bmatrix} = \begin{bmatrix} (I - L_k C_k) s_k - L_k \nu_k \\ \nu_k \end{bmatrix} = \begin{bmatrix} I - L_k C_k & -L_k \\ 0 & I \end{bmatrix} \begin{bmatrix} s_k \\ \nu_k \end{bmatrix}. \tag{14}$$

Similarly to the above proof process, the JPDF of e_k can be obtained as described by (10). This completes the proof.

Remark 2. Theorem 1 presents the JPDFs of the prediction error and the estimation error, which are denoted as $\gamma_{s_k}(\eta)$ and $\gamma_{e_k}(\tau)$ respectively. It is noted that $\gamma_{e_k}(\tau)$ is actually a conditional JPDF related to the undetermined filter gain L_k, and thus can be further expressed as $\gamma_{e_k}(\tau|L_k)$. It means that the stochastic property of the estimation error relies on the selection of the filter gain. Thus, the result given in Theorem 1 plays a key role in the filtering algorithm design.

3.2 Minimum entropy filter design

Following the formulation of the JPDF, the entropy of estimation error can be calculated. Compared with the above mentioned Shannon entropy, Renyi's entropy has the advantage of computational efficiency [22, 23]. The α-order Renyi's entropy of e_k is defined as follows [24]:

$$H_\alpha(e_k) = \frac{1}{1-\alpha} \ln \int_{[a_k,b_k]^n} \gamma_{e_k}^{\alpha}(\tau) \mathrm{d}\tau. \tag{15}$$

In (15), if we select $\alpha = 2$, it becomes the well-known quadratic Renyi's entropy,

$$H_2(e_k) = -\ln \int_{[a_k,b_k]^n} \gamma_{e_k}^{2}(\tau) \mathrm{d}\tau = -\ln V_2(e_k), \tag{16}$$

where $V_2(e_k) = \int_{[a_k,b_k]^n} \gamma_{e_k}^2(\tau)\mathrm{d}\tau$ is named quadratic information potential. From (16), it is easy to observe that the minimization of $H_2(e_k)$ is equivalent to the maximization of $V_2(e_k)$. Therefore, in order to reduce the computational complexity, the quadratic information potential changed of sign will be employed instead of entropy as the criterion for the filter design.

By minimizing $-V_2(e_k)$, the optimal filter gain L_k can be solved. When the PDF expressions of noises and initial state are relatively simple, the explicit analytic expression of the optimal filter gain may be obtained. However, for the general case, it is indeed difficult to solve this optimization problem analytically. Similarly to [10, 25], the gradient optimization technique can be adopted in this paper. In order to use this technique, a stretched column vector l_k is defined by

$$l_k = \begin{bmatrix} L_{k1} & \cdots & L_{kn} \end{bmatrix}^{\mathrm{T}}, \tag{17}$$

where $L_{kj} (j = 1, \ldots, n)$ is the jth row of filter gain matrix L_k. Furthermore, the following performance index J_k is considered in this subsection:

$$\begin{aligned} J_k &= -R_1 V_2(e_k) + \frac{1}{2} l_k^T R_2 l_k \\ &= -R_1 \int_{[a_k,b_k]^n} \gamma_{e_k}^2(\tau)\mathrm{d}\tau + \frac{1}{2} l_k^T R_2 l_k, \end{aligned} \tag{18}$$

where $R_1 > 0$ and $R_2 \geqslant 0$ are two pre-specified weights. In (18), the first term is the negative of quadratic information potential of the estimation error, and the second term reflects the constraint on the filter gain.

Remark 3. Since the estimation is unbiased naturally, there is no need to consider the mean term in the above performance index. Even for the system with noises having non-zero means, unbiasedness can also be guaranteed by adding the means of noises in the filter equations (2) and (3). Thus, performance index (18) considers the entropy of estimation error but no the mean value, which is different from the discussion in [11].

Once the performance index is established, the optimal filter gain can be obtained by using the following gradient algorithm:

$$l_k^{i+1} = l_k^i - \lambda_k \left.\frac{\partial J_k}{\partial l_k}\right|_{l_k=l_k^i} = l_k^i - \lambda_k \nabla J_k(l_k^i), \tag{19}$$

where i represents the ith iteration, $\lambda_k > 0$ is a pre-specified step length for the optimization. In this way, l_k^i can be convergent to l_k^* iteratively. Setting the initial value l_k^0 and the accuracy $\varepsilon > 0$, we could carry out the above algorithm until $|\nabla J(l_k^i)| < \varepsilon$ holds. Then, l_k^* can be approximated by l_k^i in the last iteration. To sum up, we have the following result.

Theorem 2. Under the performance index (18), the optimal filter gain matrix L_k is given by

$$L_k = \begin{bmatrix} l_{k1}^* & \cdots & l_{kn}^* \end{bmatrix}^{\mathrm{T}}, \tag{20}$$

where $l_{kj}^* \in \mathbb{R}^m$ $(j = 1, \dots, n)$ represents a sub-vector including the $[(j-1)m+1]$th–$[jm]$th elements of l_k^*, and the vector l_k^* is provided by the above gradient optimization technique.

Remark 4. From Theorems 1 and 2, it can be seen that the calculation of error PDFs and optimal filter gain does not depend on the measurement output, but depends only on the system matrices and the PDFs of noises. It means that the optimal filter gain L_k can be obtained offline and the filtering performance can also be evaluated before the system operates, which is an important practical aspect of the proposed filter.

Remark 5. Similar to the form of the classic Kalman filer, the proposed minimum entropy filter can also be summarized as the following five formulas:

- Time update (predict): (2), (9);
- Measurement update (correct): (20), (3), (10).

These results, although not concise enough in form, are more applicable to the filtering for general non-Gaussian stochastic systems.

3.3 Relative entropy based filter design

As mentioned above, the filter design objective should be to minimize the randomness of the estimation error or to realize a narrowly distributed Gaussian error signal. The former is a minimum entropy issue which has been presented in the previous subsection, and the latter is in fact a PDF shape control problem [20]. As the random noises involved in the system are of non-Gaussian type, the estimation error is also a non-Gaussian signal. In this subsection, another type of filtering algorithm is proposed where our task is to make the PDF of the estimation error follow closely a target distribution (generally a narrowly distributed Gaussian PDF with zero mean). It means that the estimation error signal is guaranteed to have the desired stochastic property by minimizing the distance between the error distribution and the target one.

In order to measure the distance between the estimation error PDF and the target one, the relative entropy (also named Kullback-Leibler divergence) is employed in this subsection, which is defined as follows [12, 26, 27]:

$$D_{\mathrm{KL}}(\gamma_{e_k}(\tau)\|\gamma_\phi(\tau)) = \int_{[a_k,b_k]^n} \gamma_{e_k}(\tau) \ln \frac{\gamma_{e_k}(\tau)}{\gamma_\phi(\tau)} \mathrm{d}\tau, \tag{21}$$

where $\gamma_\phi(\tau)$ is a pre-specified JPDF for the error JPDF $\gamma_{e_k}(\tau)$ to follow. The relative entropy is known to be non-negative. It is zero for $\gamma_{e_k}(\tau) = \gamma_\phi(\tau)$ almost everywhere, and infinite if there exists a set with a positive Lebesgue measure on which $\gamma_\phi(\tau) \equiv 0$ and $\gamma_{e_k}(\tau) > 0$.

Although it is an effective tool to measure the distance between two distributions, the relative entropy is not actually a true metric or distance because it does not satisfy the symmetry or the triangle inequality [28]. Since we only focus on the optimization problem for the filter design, the exact quantity of divergence does not matter. Thus, we still adopt the relative entropy (21) for the sake of simplicity.

By minimizing the relative entropy, the JPDF of the estimation error can be made as close as possible to the target $\gamma_\phi(\tau)$. Therefore, the corresponding filter is named as relative entropy based filter or PDF tracking filter.

Similar to (18), the performance index for the PDF tracking filter is formulated as

$$\bar{J}_k = R_3 D_{\mathrm{KL}}(\gamma_{e_k}(\tau)\|\gamma_\phi(\tau)) + \frac{1}{2}l_k^T R_4 l_k$$
$$= \int_{[a_k,b_k]^n} R_3\gamma_{e_k}(\tau)\ln\frac{\gamma_{e_k}(\tau)}{\gamma_\phi(\tau)}\mathrm{d}\tau + \frac{1}{2}l_k^T R_4 l_k, \quad (22)$$

where $R_3 > 0$ and $R_4 \geqslant 0$ are two weights. If the target PDF $\gamma_\phi(\tau)$ is selected as a Gaussian PDF, the relative entropy is naturally bounded simply because the Lebesgue measure of the set on which Gaussian function takes zero value is just equal to zero. In this case, the weight R_3 is taken to be a positive constant. Otherwise, for more general case, a weighting function $R_3(\tau)$ should be designed to guarantee the boundedness of the first term in performance index (22). The similar problem has been discussed in [12], and the selection of $R_3(\tau)$ can refer to it.

Furthermore, the optimal filter gain can be obtained by minimizing the performance index $\bar{J}_k$. The specific solution algorithm is just similar to the one presented in the previous subsection, which is omitted here for brevity.

Remark 6. To sum up, two kinds of filtering algorithms including minimum entropy filter and relative entropy based filter (PDF tracking filter) have both been designed for the non-Gaussian stochastic system (1). Different from the classic Kalman filter, the entropy optimization based filters are more applicable for general non-Gaussian systems. Compared with the previous studies [11, 12], this paper focuses on a wider class of systems and fully analyzes the properties of priori and posteriori estimates. Therefore, the results provided in this paper are the extensions on previous work.

4 Illustrative examples

4.1 Minimum entropy filtering

To demonstrate the proposed minimum entropy filtering algorithm, we consider a simple model described by

$$\begin{cases} x_k = (0.7 - 0.1\tan(1+k)^{-1})x_{k-1} + \omega_{k-1}, \\ z_k = 0.8x_k + \nu_k, \end{cases}$$

with the initial state x_0 assumed to obey

$$\gamma_{x_0}(\tau) = \begin{cases} -\frac{3}{32}(\tau-3)(\tau-7), & \tau \in [3,7], \\ 0, & \text{otherwise.} \end{cases}$$

Random noises ω_k and ν_k are assumed to be mutually independent, where ω_k is subjected to an uniform distribution defined on the interval $[-0.2, 0.2]$ and the PDF of ν_k is denoted by

$$\gamma_{\nu_k}(\tau) = \begin{cases} -6(\tau^2 - 0.25), & \tau \in [-0.5, 0.5], \\ 0, & \text{otherwise.} \end{cases}$$

Besides, both ω_k and ν_k are independent of x_0.

The minimum entropy filter is designed according to Remark 5, with the initial value set to be $\hat{x}_0^+ = 5$. The weights in performance index (18) are selected as $R_1 = 1$ and $R_2 = 0$. At each sample time, gradient algorithm (19) is applied, where the initial condition is taken to be $l_k^0 = 0.6$. The simulation results are displayed in Figures 1–5.

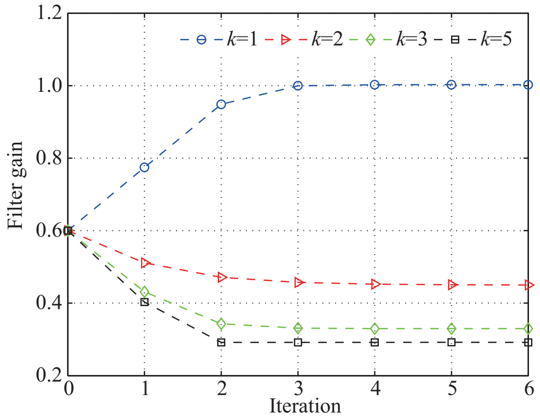

Figure 1 (Color online) Filter gain iteration at some sample times.

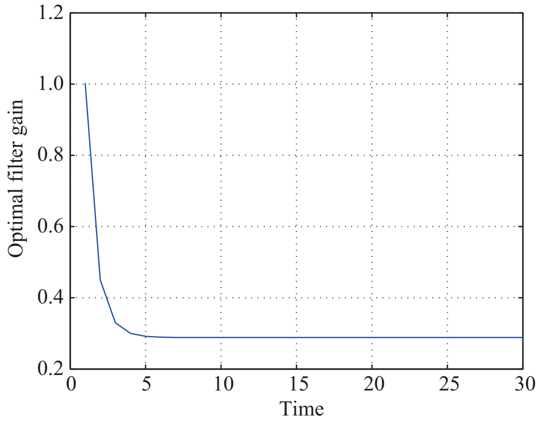

Figure 2 (Color online) Optimal filter gain.

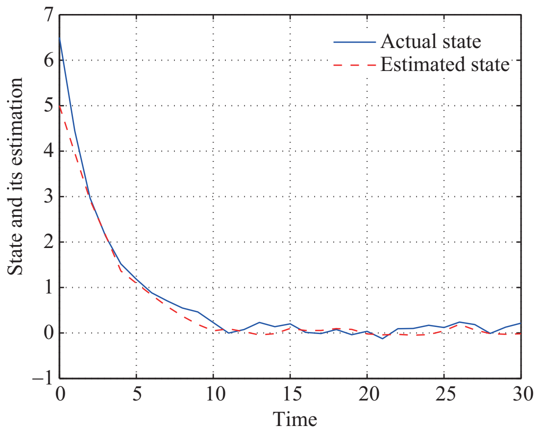

Figure 3 (Color online) System state and its estimation.

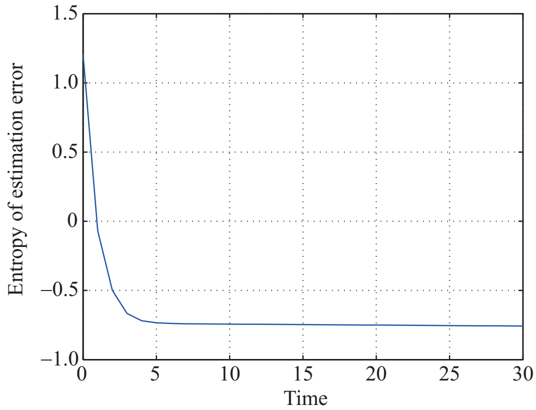

Figure 4 (Color online) Entropy of the estimation error.

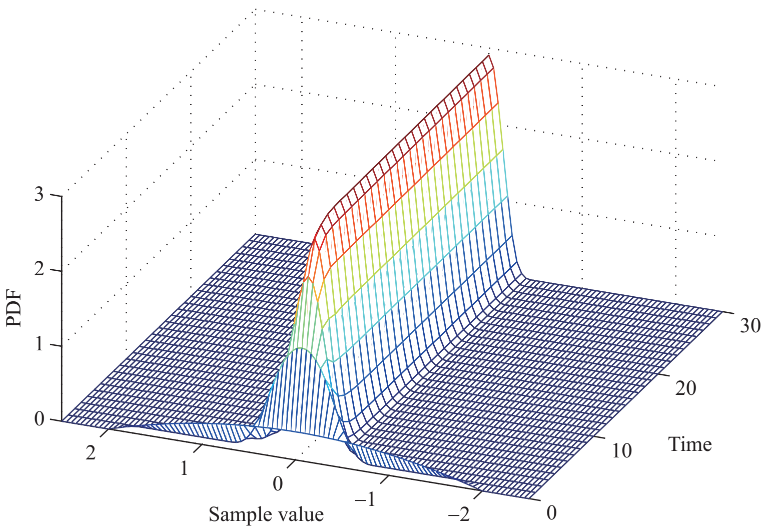

Figure 5 (Color online) PDF of the estimation error.

Figure 1 shows the solution processes of filter gain at several sample times ($k = 1, 2, 3, 5$). It can be seen that at each time the filter gain is convergent to the optimal value under the gradient algorithm.

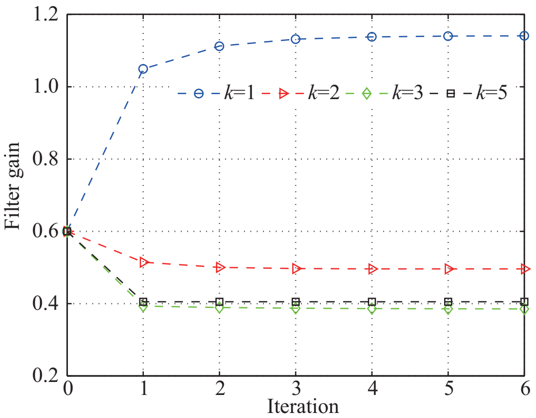

Figure 6 (Color online) Filter gain iteration at some sample times.

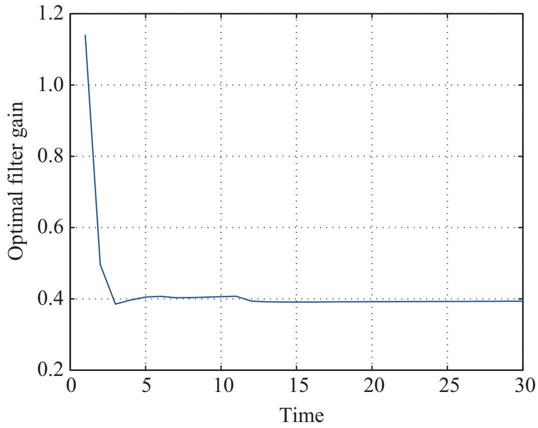

Figure 7 (Color online) Optimal filter gain.

Figure 2 shows the sequence of the obtained optimal filter gain. The dynamical responses of system state x_k and estimated state $\hat{x}_k^+$ are displayed in Figure 3, which shows that the estimated state can be able to follow the actual state closely. The entropy and PDF curves of estimation error are presented in Figures 4 and 5, respectively. It can be seen that the estimation error has narrower PDF curve corresponding to less randomness. All these simulation results demonstrate the validity of the proposed minimum entropy filtering algorithm.

4.2 Relative entropy based filtering

In this subsection, the relative entropy based filtering (PDF tracking filtering) algorithm is tested. The system model in the previous subsection is still used, with the initial state assumed to obey an uniform distribution on the interval $[3, 7]$. The distribution of the process noise ω_k is the same as above, and the measurement noise ν_k has a nonsymmetric PDF defined by

$$\gamma_{\nu_k}(\tau) = \begin{cases} -12(t + \frac{24}{35})^5 + 12(t + \frac{24}{35})^3, & \tau \in [-\frac{24}{35}, \frac{11}{35}], \\ 0, & \text{otherwise.} \end{cases}$$

The desired PDF $\gamma_\phi(\tau)$ is supposed to be a Gaussian distribution $N(0, 0.12^2)$.

In the simulation, the initial estimate is $\hat{x}_0^+ = 5$, the initial value of the filter gain iteration at each sample time is taken to be $l_k^0 = 0.6$, and the weights in performance index (22) are set to be $R_3 = 1$ and $R_4 = 0$. The corresponding simulation results are displayed in Figures 6–10.

Similar to Figures 1 and 2, Figures 6 and 7 show the solution processes of optimal filter gain at several sample times and the sequence of obtained optimal gain, respectively. Figure 8 displays the time responses of actual state and estimated state. In Figure 9, the PDFs of estimation error at two sample times ($k = 1, 30$) as well as the target PDF $\gamma_\phi(\tau)$ are given. Moreover, the 3D mesh plot of the error PDF is displayed in Figure 10. It can be seen that the distribution of estimation error can follow the target distribution closely. In other words, the error PDF has been made as narrow and as Gaussian as possible under the designed PDF tracking filtering algorithm. From these results, we arrive at the conclusion that the proposed algorithm has a favorable performance.

5 Conclusion

The filtering problem for multivariate stochastic systems with simultaneous presence of non-Gaussian process noise and measurement noise is investigated in this paper. The form of the filter consists of time update and measurement update two steps, and the filter gain in the measurement update equation is

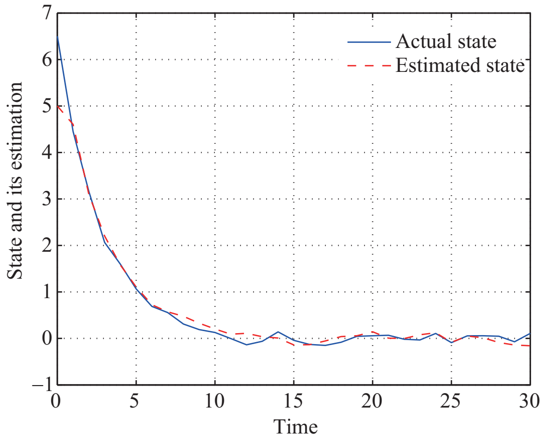

Figure 8 (Color online) System state and its estimation.

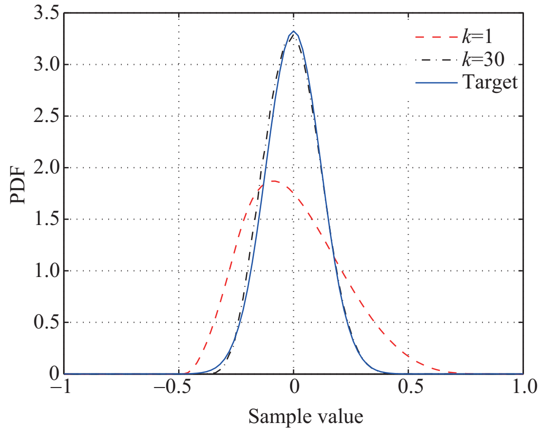

Figure 9 (Color online) Comparisons between the error PDF and the target.

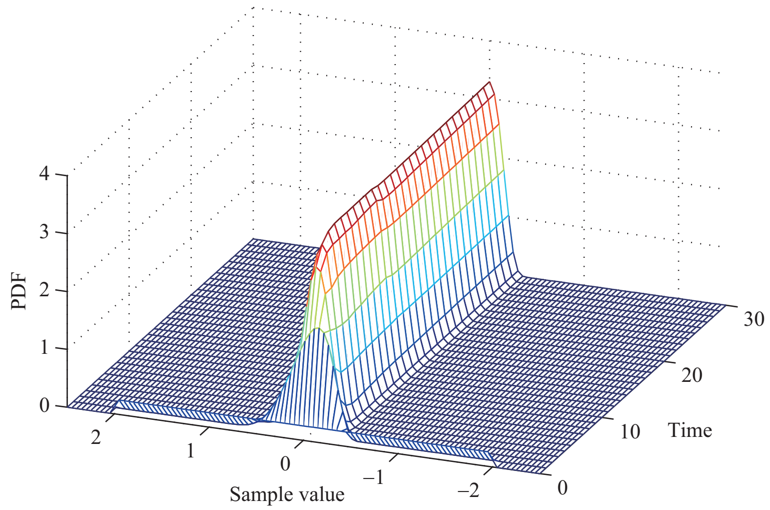

Figure 10 (Color online) PDF of the estimation error.

designed based on the principle that the randomness of the estimation error should be minimized. For this purpose, two concepts including quadratic information potential and relative entropy are employed as the performance criteria respectively. The relationship between the JPDFs of noises and estimation error is formulated, based on which the two performance indices are established. Following the solution of optimal filter gain, two kinds of filtering algorithms are obtained, and named minimum entropy filter and relative entropy based filter respectively. Finally, numerical simulation examples are given to verify the effectiveness of the proposed algorithms. However, some issues still need to be studied in our future work. Examples include the the further improvement of the filter performance and the extension to nonlinear stochastic systems.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61320106010, 61573019, 61627810). The authors would like to thank the anonymous reviewers for their constructive comments which helped to improve the quality and presentation of this paper significantly.

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Kalman R E. A new approach to linear filtering and prediction problems. Trans ASME J Basic Eng, 1960, 82: 35–45

2 Anderson B D, Moore J B. Optimal Filtering. Englewood Cliffs: Prentice Hall, 1979

3 Simon D. Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches. New York: Wiley, 2006

4 Ding D R, Wang Z D, Ho D W C, et al. Distributed recursive filtering for stochastic systems under uniform quantizations and deception attacks through sensor networks. Automatica, 2017, 78: 231–240

5 Ding D R, Wei G L, Zhang S J, et al. On scheduling of deception attacks for discrete-time networked systems equipped with attack detectors. Neurocomputing, 2017, 219: 99–106

6 Fridman E, Shaked U. An improved delay-dependent H_∞ filtering of linear neutral systems. IEEE Trans Signal Process, 2004, 52: 668–673

7 Wang Z D, Liu Y R, Liu X H. H_∞ filtering for uncertain stochastic time-delay systems with sector-bounded nonlinearities. Automatica, 2008, 44: 1268–1277

8 Yang F W, Li Y M, Liu X H. Robust error square constrained filter design for systems with non-Gaussian noises. IEEE Signal Process Lett, 2008, 15: 930–933

9 Wei G L, Wang Z D, Shen B. Error-constrained filtering for a class of nonlinear time-varying delay systems with non-Gaussian noises. IEEE Trans Autom Control, 2010, 55: 2876–2882

10 Yue H, Wang H. Minimum entropy control of closed-loop tracking error for dynamic stochastic systems. IEEE Trans Autom Control, 2003, 48: 118–122

11 Guo L, Wang H. Minimum entropy filtering for multivariate stochastic systems with non-Gaussian noises. IEEE Trans Autom Control, 2006, 51: 695–700

12 Zhou J L, Zhou D H, Wang H, et al. Distribution function tracking filter design using hybrid characteristic functions. Automatica, 2010, 46: 101–109

13 Saridis G N. Entropy formulation of optimal and adaptive control. IEEE Trans Autom Control, 1988, 33: 713–721

14 Petersen I R, James M R, Dupuis P. Minimax optimal control of stochastc uncertain systems with relative entropy constraint. IEEE Trans Autom Control, 2000, 45: 398–412

15 Zhang J H, Chu C C, Munoz J, et al. Minimum entropy based run-to-run control for semiconductor processes with uncertain metrology delay. J Process Control, 2009, 19: 1688–1697

16 Yi Y, Guo L, Wang H. Adaptive statistic tracking control based on two steps neural networks with time delays. IEEE Trans Neural Netw, 2009, 20: 420–429

17 Ding J L, Chai T Y, Wang H. Offline modeling for product quality prediction of mineral processing using modeling error PDF shaping and entropy minimization. IEEE Trans Neural Netw, 2011, 22: 408–419

18 Liu Y L, Wang H, Guo L. Observer-based feedback controller design for a class of stochastic systems with non-Gaussian variables. IEEE Trans Autom Control, 2015, 60: 1445–1450

19 Wang H. Bounded Dynamic Stochastic Systems: Modeling and Control. London: Springer-Verlag, 2000

20 Guo L, Wang H. Stochastic Distribution Control System Design: a Convex Optimization Approach. London: Springer-Verlag, 2010

21 Papoulis A. Probability, Random Variables and Stochastic Processes. 3rd ed. New York: McGraw-Hill, 1991

22 Guo L, Yin L P, Wang H, et al. Entropy optimization filtering for fault isolation of nonlinear non-Gaussian stochastic systems. IEEE Trans Autom Control, 2009, 54: 804–810

23 Afshar P, Wang H, Chai T Y. An ILC-based adaptive control for general stochastic systems with strictly decreasing entropy. IEEE Trans Neural Netw, 2009, 20: 471–482

24 Renyi A. A Diary on Information Theory. New York: Wiley, 1987

25 Ren M F, Zhang J H, Wang H. Minimized tracking error randomness control for nonlinear multivariate and non-Gaussian systems using the generalized density evolution equation. IEEE Trans Autom Control, 2014, 59: 2486–2490

26 Karny M. Toward fully probabilistic control design. Automatica, 1996, 32: 1719–1722

27 Yin L P, Zhang H Y, Guo L. Data driven output joint probability density function control for multivariate non-linear non-Gaussian systems. IET Control Theory Appl, 2015, 9: 2697–2703

28 Liu Y, Wang H, Hou C H. Sliding-mode control design for nonlinear systems using probability density function shaping. IEEE Trans Neural Netw Learn Syst, 2014, 25: 332–343