entropy

Article

Improved Minimum Entropy Filtering for Continuous Nonlinear Non-Gaussian Systems Using a Generalized Density Evolution Equation

Mifeng Ren [1], Jianhua Zhang [2,*], Fang Fang [1], Guolian Hou [1] and Jinliang Xu [3]

[1] School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; E-Mails: renmifeng@126.com (M.R.); ffang@ncepu.edu.cn (F.F.); hgl@ncepu.edu.cn (G.H.)

[2] State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China

[3] The Beijing Key Laboratory of New and Renewable Energy, North China Electric Power University, Beijing 102206, China; E-Mail: xjl@ncepu.edu.cn

* Author to whom correspondence should be addressed; E-Mail: zjh@ncepu.edu.cn; Tel.: +86-10-6177-2106.

Abstract: This paper investigates the filtering problem for multivariate continuous nonlinear non-Gaussian systems based on an improved minimum error entropy (MEE) criterion. The system is described by a set of nonlinear continuous equations with non-Gaussian system noises and measurement noises. The recently developed generalized density evolution equation is utilized to formulate the joint probability density function (PDF) of the estimation errors. Combining the entropy of the estimation error with the mean squared error, a novel performance index is constructed to ensure the estimation error not only has small uncertainty but also approaches to zero. According to the conjugate gradient method, the optimal filter gain matrix is then obtained by minimizing the improved minimum error entropy criterion. In addition, the condition is proposed to guarantee that the estimation error dynamics is exponentially bounded in the mean square sense. Finally, the comparative simulation results are presented to show that the proposed MEE filter is superior to nonlinear unscented Kalman filter (UKF).

Keywords: non-Gaussian systems; stochastic filtering; generalized density evolution equation; exponentially mean-square boundedness

1. Introduction

State estimation theory has been regarded as an important research area in modern control systems. In particularly, the appearance of the Kalman filtering theory in the last century had a profound influence on modern optimal control [1–3]. Kalman filter algorithm is based on the minimum variance estimation for the linear Gaussian systems. However, the noises in real systems may be non-Gaussian; and even for a nonlinear system with Gaussian noises, the system output could be a non-Gaussian variable. In these cases, Kalman filter may lead to poor estimation.

Some attempts have been made at studying filtering algorithms for nonlinear systems with non-Gaussian noises. The existed methodologies to design filters for nonlinear non-Gaussian systems can be classified into three kinds: analytical approach, simulation-based approach and adaptive approach.

Analytical approaches to filtering for nonlinear non-Gaussian systems were investigated in [4–7]. In [4,5], nonlinear non-Gaussian systems were described by combining an improved square-root B-spline model with a further nonlinear dynamic model. Once B-spline expansions have been made for probability density functions (PDFs), further modeling was still needed to reveal the relationship between the input and the weights related to the PDFs, a nonlinear filter was then constructed by minimizing the error between the measured output PDF and estimated output PDF. Although the proposed filter is suitable to nonlinear non-Gaussian systems, the output PDF should be measureable. Moreover, it is not easy to build the state space expression of the weights related to the PDFs in practical systems. The filter for a class of multivariate dynamic stochastic systems with non-Gaussian stochastic input and nonlinear output was investigated in [6] and [7] respectively. In [6], a new formulation of the residual PDF was made to link the residual PDF to the gain matrix of the filter, and the optimal filtering gain matrix was then solved by minimizing the entropy of the residual. The minimum entropy filter in [6] presented a good performance in reducing the randomness of the filter residual, and was more general and suited for non-Gaussian systems. However, minimum entropy criterion may not guarantee that the estimation errors approach to zero. In [7], following the minimum information divergence criterion, a hybrid characteristic function of the conditional estimation error was introduced to construct the performance index of the tracking filter. An analytical solution of the filter gain matrix was then obtained so that the PDFs of the filtering error can follow a target distribution shape. Nevertheless, it is a little complicated to calculate the analytical solution.

There are two types of simulation-based approaches to filtering for nonlinear non-Gaussian systems: numerical integration and sequential Monte Carlo. Although the filtering problem can be tackled using numerical integration [8,9], it is difficult to implement when the dimension of the state vector is higher. Sequential Monte Carlo simulation [10–15], which is also named particle filtering strategy, has shown its great advantages to deal with filtering problems for nonlinear non-Gaussian systems. Nevertheless, there are still many issues to be solved, for example, (1) random sampling may bring the accumulation of Monte Carlo error and even lead to filter divergence; (2) a large number of particles are needed to avoid degradation and to improve the estimation accuracy, which makes the calculation a sharp increase.

Adaptive approaches have been also investigated for solving filtering problem in nonlinear non-Gaussian systems in last decades. Since minimum error entropy criterion ensures that the estimation error has small uncertainty, it was used for supervised training of nonlinear stochastic

system in [16,17]. However, entropy does not change with the mean of the distribution, the algorithm may not yield zero-mean error. Therefore, the result may be corrected by properly modifying the bias of the output processing element of the neural networks. Maximum mutual information criterion was proposed for adaptive filtering in [18], this criterion is robust to measure distortions. Nevertheless, the maximum mutual information criterion leads to non-unique optimum solution. It is necessary to use a priori information about the unknown system in order to obtain unique solution.

In general, the filtering problem for nonlinear non-Gaussian systems calls for further investigation. The filtering problem addressed here is solved by combining the improved performance index with optimal design method.

Entropy is a natural extension beyond mean square error because it is a function of PDF. One of the most important problems for minimum entropy filtering is the formulation of the PDF of estimation error. For continuous nonlinear systems, the classical Liouville equation and its Dostupov-Pugachev extension have been used to obtain the PDF of the concerned stochastic variable. However, it is very difficult to have their analytical/numerical solutions due to high-dimensional partial differential equations. From the viewpoint of the random event description of the principle of preservation of probability, the generalized density evolution equation was developed in [19]. Different from the traditional density evolution equations, the dimensions of the generalized density evolution equation just depend on the necessary physical quantities' dimensions not the whole original systems itself, which makes the dimensions of the partial differential equations much lower. Some application results (see e.g., [20,21]) illustrated the efficiency and conveniences of the formulated density evolution equation.

The contribution of this paper is to develop a new filtering strategy for multivariable nonlinear systems with non-Gaussian disturbances by utilizing a novel performance index which contains the entropy of estimation error, square error and constraints on gain matrix of the filter. A novel approach, which uses the principle of preservation of probability, is presented here to formulate the joint PDF evolution equation of the estimation errors. The PDF evolution equation explicitly reveals the relationship among the estimation errors of the filter, filter gain and random inputs. In addition, the entropy of estimation errors in the performance index is replaced by its information potential presented in [16,17] so as to simplify the calculation of the entropy. The filter gain matrix for nonlinear non-Gaussian systems is then designed by minimizing the proposed performance index. This filtering algorithm yields to the estimation errors which not only have small uncertainty but also approach to zero. Finally, the exponentially boundedness in the mean square sense is analyzed for the estimation error dynamics.

This paper is organized as follows: in section 2, the state-space model of a nonlinear non-Gaussian system and the structure of filter are built to formulate the filtering problem. The joint PDF evolution of estimation errors is formulated in Section 3 so as to calculate the entropy of estimation errors. The filtering algorithm is proposed in Section 4 by minimizing the improved entropy criterion and the exponentially mean-square boundedness condition for error systems is provided. Comparative simulation results are given to illustrate the efficiency and validity of the proposed method in Section 5. Conclusions are drawn in Section 6. *Notation*. $\Re^n$ and $\Re^{n \times m}$ denote the n-dimensional Euclidean space and the set of all $n \times m$ real matrices respectively. The superscript "T" denotes the transpose. If A is a

matrix, $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ represent the largest and smallest eigenvalue of A respectively. $\mathrm{E}(\cdot)$ stands for the mathematical expectation of random variables.

2. Problem Formulation

2.1. System Model

Consider a nonlinear stochastic continuous system described by:

$$\begin{cases} \dot{x}(t) = A(t)x(t) + D(t)\omega(t) \\ y(t) = F(x(t), \upsilon(t)) \end{cases} \tag{1}$$

where $x \in \Re^n$ is the state vector, $y \in \Re^m$ is the measured output vector. $\omega \in \Re^p$ and $\upsilon \in \Re^q$ are system noises and measurement noises respectively. $A(t)$ and $D(t)$ are two known time-varying system matrices. The following assumptions are made for simplicity, which can be satisfied by many practical systems.

Assumption 1: $F(\cdot)$ is a known Borel measurable and smooth vector-value nonlinear function of its arguments.

Assumption 2: ω and υ are bounded, mutually independent random vectors with known joint PDFs γ_ω and γ_υ.

2.2. Filter Dynamics

For the nonlinear stochastic system (1), the filter dynamics can be formulated as follows:

$$\begin{cases} \dot{\hat{x}}(t) = A(t)\hat{x}(t) + L(y(t) - \hat{y}(t)) \\ \hat{y}(t) = F(\hat{x}(t), 0) \end{cases} \tag{2}$$

The resulting estimation error $e = x - \hat{x}$ satisfies:

$$\dot{e}(t) = \dot{x}(t) - \dot{\hat{x}}(t) = A(t)e(t) - L\left(F(x, \upsilon) - F(\hat{x}, 0)\right) + D(t)\omega(t) \tag{3}$$

where $L \in \Re^{n \times m}$ is the gain matrix to be determined and can be denoted as $L = \begin{bmatrix} L_1^T & \cdots & L_n^T \end{bmatrix}^T$, L_i is the *ith* row vector of L. Let $l = \begin{bmatrix} L_1 & \cdots & L_n \end{bmatrix}^T$ and thus $l \in \Re^{mn \times 1}$ is a stretched column vector.

The purpose of filtering algorithm for nonlinear non-Gaussian systems is to ensure that the estimation errors achieve small dispersion and approach to zero. Entropy provides a measure of the average information contained in a given PDF. When entropy of estimation errors is minimized, all moments of the error PDF are constrained. Hence, the following quadratic Renyi's entropy of estimation error is still contained in the performance index to measure the dispersion of the estimation error:

$$H_2(e) = -\log \int_{\Omega_e} p_e^2(z)dz = -\log V(e) \tag{4}$$

where $p_e(z)$ is the joint PDF of estimation error and $V(e)$ the information potential [14]. It can be seen that it is necessary to provide the joint PDF of errors firstly. In the next section, we will deduce the expression of $p_e(z)$ according to the generalized density evolution equation.

3. Formulation for the Joint PDF of Error

In most cases, Equation (3) is a well-posed equation, and the error vector $e(t)$ can be determined uniquely. It may be assumed to be:

$$e(t) = H(\omega, \upsilon, L, t) \tag{5}$$

At the present stage, the explicit expression of $H(\cdot)$ is not requisite, and the sufficient condition just needs to know its existence and uniqueness. The derivative of $e(t)$ can be assumed to take the form:

$$\dot{e}(t) = h(\omega, \upsilon, L, t) \tag{6}$$

It is observed that all the randomness involved in this error dynamics comes from noises ω and υ, thus, the augmented system $(e(t), \omega, \upsilon)$ is probability preserved. In other words, from the random event description of the principle of preservation of probability, it leads to:

$$\frac{D}{Dt} \int_{\Omega_e \times \Omega_\omega \times \Omega_\upsilon} p_{e\omega\upsilon}(e, \omega, \upsilon, t) de d\omega d\upsilon = 0 \tag{7}$$

where Ω_e, Ω_ω and Ω_υ are the distribution domains of e, ω and υ, respectively; $p_{e,\omega,\upsilon}(e, \omega, \upsilon, t)$ is the joint PDF of $(e(t), \omega, \upsilon)$. It follows from Equation (7) that:

$$\frac{D}{Dt} \int_{\Omega_e \times \Omega_\omega \times \Omega_\upsilon} p_{e\omega\upsilon}(e, \omega, \upsilon, t) de d\omega d\upsilon = \int_{\Omega_e \times \Omega_\omega \times \Omega_\upsilon} \left(\frac{\partial p_{e\omega\upsilon}}{\partial t} + h \frac{\partial p_{e\omega\upsilon}}{\partial e} \right) de d\omega d\upsilon \tag{8}$$

Combining Equation (7) with Equation (8) and considering the arbitrariness of $\Omega_e \times \Omega_\omega \times \Omega_\upsilon$, it yields:

$$\frac{\partial p_{e\omega\upsilon}(e, \omega, \upsilon, t)}{\partial t} + \dot{e}(\omega, \upsilon, t) \frac{\partial p_{e\omega\upsilon}(e, \omega, \upsilon, t)}{\partial e} = 0 \tag{9}$$

where Equation (6) and Equation (9) are the generalized density evolution equation (GDEE) for $(e(t), \omega, \upsilon)$. The corresponding instantaneous PDF of $e(t)$ can be obtained by solving a family of partial differential equations with the following given initial condition:

$$p_{e\omega\upsilon}(e, \omega, \upsilon, t)\big|_{t=0} = \delta(e - e_0) p_\omega(\omega) p_\upsilon(\upsilon) \tag{10}$$

where $\delta(\cdot)$ is the Dirac-Delta function; and $e_0 = x_0 - \hat{x}_0$ is deterministic initial value of $e(t)$. Then, we have:

$$p_e(e, t) = \int_{\Omega_\omega \times \Omega_\omega} p_{e\omega\upsilon}(e, \omega, \upsilon, t) d\omega d\upsilon \tag{11}$$

where the joint PDF $p_{e\omega\upsilon}(e, \omega, \upsilon, t)$ is the solution of (8), which can be obtained according to the method presented in [22].

4. Improved MEE Filtering

4.1. Performance Index

The following performance index for a minimum entropy filter was presented in [6]:

$$J_1 = -R_1 \int_{\Omega_e} p(e,t) \log p(e,t) de + \frac{1}{2} l^T R_2 l \tag{12}$$

where R_1 and R_2 are weighing matrices. The first term is the entropy of the estimation errors and the second term is used to penalize the elements of the gain. Although the MEE criteria can minimize both the probabilistic uncertainty and dispersion of the estimation error, it cannot guarantee the minimum error. In this work , an improved performance index J is considered:

$$J = R_1 H_2(e(t)) + \frac{1}{2} l^T(t) R_2 l(t) + R_3 E(e^T(t)e(t)) \tag{13}$$

where $H_2(e(t))$ is the quadratic Renyi's entropy given in Equation (4). $R_3 > 0$ is the weight corresponding to the mean squared error. The third term on the right side of the equation is utilized to make the estimation errors approach to zero. Since Renyi's entropy is a monotonic increasing function of the negative information potential, minimization of Renyi's entropy is equivalent to minimizing the inverse of the quadratic information potential $\dfrac{1}{V(e)}$, so the performance index can be rewritten as follows:

$$\begin{aligned}
J &= \frac{R_1}{V(e(t))} + \frac{1}{2} R_2 l^T(t) l(t) + R_3 E(e^T(t)e(t)) \\
&= \frac{R_1}{\int_{\Omega_e} p_e^2(\tau,l,t) d\tau} + \frac{1}{2} R_2 l^T(t) l(t) + R_3 \int_{\Omega_e} e^T(t) e(t) p_e(\tau,l,t)) d\tau
\end{aligned} \tag{14}$$

Remark 1: The entropy of estimation error is replaced by quadratic information potential of error in order to simplify the calculation. Nonparametric estimation of the PDF of estimation errors is another alternative way to estimate the PDF except the method introduced in [22]. Nonparametric entropy (or information potential) estimation of random error can be found in [16,17].

4.2. Optimal Filter Gain Matrix

After constructing the performance index J , the optimal filter gain can be carried out by minimizing the performance index Equation (14), which can be regarded as an unconstrained nonlinear programming problem. The classical conjugate gradient approach is employed here to obtain an iterative solution, which could avoid the difficulty of solving the extremum condition ($\frac{\partial J}{\partial l} = 0$). The elements of optimal filter gain can be solved and summarized in Theorem 1.

Theorem 1: For a given accuracy $\varepsilon > 0$, the gain matrix of the suboptimal improved MEE filter is given by:

$$l_{k+1} = l_k + \eta_k P_k \tag{15}$$

where $\eta_k = \dfrac{-\nabla J(l_k)^T P_k}{P_k^T H(l_k) P_k}, P_k = \begin{cases} -\nabla J(l_k), & k = 0 \\ -\nabla J(l_k) + \beta_{k-1} P_{k-1}, & k > 0 \end{cases}$ and the Hessian matrix $H(l_k)$ is:

$$H(l_k) = \left[\frac{\partial^2 J(l_k)}{\partial l_i \partial l_j}\right]_{mn \times mn} = \begin{bmatrix} \dfrac{\partial^2 J(l_k)}{\partial l_1^2} & \cdots & \dfrac{\partial^2 J(l_k)}{\partial l_1 \partial l_{mn}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 J(l_k)}{\partial l_{mn} \partial l_1} & \cdots & \dfrac{\partial^2 J(l_k)}{\partial l_{mn}^2} \end{bmatrix} \tag{16}$$

Proof: Choose an initial value l_0, and denote $k := 0$. Search the next value l_{k+1} from the given point l_k along the conjugate gradient direction P_k. Repeat the above procedure until reach the given accuracy ε. Consequently, we can obtain the optimal filtering algorithm shown in Equation (15).

Based on the conjugate gradient method, the steps to solve the optimal filter gain are summarized as follows:

Step (1) Give an accuracy $\varepsilon > 0$ and initialize l_0, denote $k = 0$

Step (2) Set $P_0 = -\nabla J(l_0)$.

Step (3) Obtain $\eta_k = \dfrac{-P_k^T \nabla J(l_k)}{P_k^T H(l_k) P_k}$ by solving $\min_\eta J(l_k + \eta P_k)$, and update the gain $l_{k+1} = l_k + \eta_k P_k$.

Step (4) If $\nabla J(l_{k+1}) < \varepsilon$, stop, and l_{k+1} is the optimal solution; Otherwise, if $k < mn - 1$, turn to step (5), and if $k = mn - 1$, reset $l_0 = l_{mn}$, and turn to step (2).

Step (5) Choose $P_{k+1} = -\nabla J(l_{k+1}) + \beta_k P_k$, set $k := k + 1$, and turn to step 3), where $\beta_k = \dfrac{\nabla f(X_{k+1})^T \nabla f(X_{k+1})}{\nabla f(X_k)^T \nabla f(X_k)}$.

Remark 2: The conjugate gradient method here could overcome the slow convergence of the steepest descent approach. It is one of the most effective algorithms to deal with large-scale nonlinear optimization problems.

4.3. Exponentially Bounded in the Mean Square

In this section, the Lyapunov method is utilized to give the condition under which the filter error dynamics is exponentially ultimately bounded in the mean square. Firstly, in order to simplify the problem, we give the following assumption:

Assumption 3: The nonlinear vector function $F(\cdot,\cdot)$ is assumed to satisfy $F(0,0) = 0$ [23]and:

$$\left\| F\big(x(t) + \sigma, \upsilon(t) + \delta\big) - F\big(x(t), \upsilon(t)\big) - \begin{bmatrix} C & B \end{bmatrix}\begin{bmatrix} \sigma \\ \delta \end{bmatrix} \right\| \le a \left\| \begin{bmatrix} \sigma \\ \delta \end{bmatrix} \right\| \tag{17}$$

where $C \in \Re^{m \times n}$, $B \in \Re^{m \times q}$ are known constant matrices, $\sigma \in \Re^n$, $\delta \in \Re^q$ are vectors, a is known positive constant.

For estimation error dynamics (3), the exponentially mean-square boundedness is defined as follows:

Definition 1 [24]: For all initial conditions $e_0 \in \Re^n$, the dynamics of the estimation error (*i.e.,* the solution of the system (3)) is exponentially ultimately bounded in the mean square if there exist constants $\alpha > 0$, $\beta > 0$ and $\gamma > 0$ such that:

$$E\left[\left\|e\left(t,t_0,e_0\right)\right\|^2\right] \le \alpha e^{-\beta t} \sup_{t_0} E\left[\left\|e_0\right\|^2\right] + \gamma \tag{18}$$

Moreover, filter (2) is said to be exponential if, for every $e_0 \in \Re^n$, system (3) is exponentially ultimately bounded in the mean square.

Theorem 2: Let the filter parameter L be given. If there exist positive scalars ε_1, ε_2 and ε_3 such that the following matrix inequality:

$$\left(A-LC\right)^T P + P\left(A-LC\right) + P\left(\varepsilon_1 LL^T + \varepsilon_2 LBB^T L^T + \varepsilon_3 DD^T\right)P + \varepsilon_1^{-1}a^2 I < 0 \tag{19}$$

has a positive definite solution $P > 0$, then system (3) is exponentially ultimately bounded in the mean square.

Proof: Fix $e_0 \in \Re^n$ arbitrarily and denote $e\left(t,t_0,e_o\right) = e\left(t\right)$. The Lyapunov function candidate can be chosen as:

$$V\left(e\left(t\right)\right) = e^T\left(t\right)Pe\left(t\right) \tag{20}$$

where $P > 0$ is the solution to the matrix inequality Equation (19).

For notational convenience, we give the following definitions:

$$\mu\left(t\right) = F\left(x,\upsilon\right) - F\left(\hat{x},0\right) - Ce\left(t\right) - B\upsilon\left(t\right) \tag{21}$$

and then it follows from Equation (3) that:

$$\dot{e}\left(t\right) = \left(A-LC\right)e\left(t\right) - L\mu\left(t\right) - LB\upsilon\left(t\right) + D\omega\left(t\right) \tag{22}$$

Hence, the derivative of V along a given trajectory Equation (22) is obtained as:

$$\begin{aligned}
\dot{V}\left(e\left(t\right)\right) = {} & e^T\left(t\right)\left[\left(A-LC\right)^T P + P\left(A-LC\right)\right]e\left(t\right) - e^T\left(t\right)PL\mu\left(t\right) \\
& - \mu^T\left(t\right)L^T Pe\left(t\right) - e^T\left(t\right)PLB\upsilon\left(t\right) - \upsilon^T\left(t\right)B^T L^T Pe\left(t\right) \\
& + e^T\left(t\right)PD\omega\left(t\right) + \omega^T\left(t\right)D^T Pe\left(t\right)
\end{aligned} \tag{23}$$

Let ε_1, ε_2 and ε_3 be positive scalars. Then the following matrix inequalities hold:

$$-e^T\left(t\right)PL\mu\left(t\right) - \mu^T\left(t\right)L^T Pe\left(t\right) \le \varepsilon_1 e^T\left(t\right)\left(PLL^T P\right)e\left(t\right) + \varepsilon_1^{-1}\mu^T\left(t\right)\mu\left(t\right) \tag{24}$$

$$-e^T\left(t\right)PLB\upsilon\left(t\right) - \upsilon^T\left(t\right)B^T L^T Pe\left(t\right) \le \varepsilon_2 e^T\left(t\right)\left(PLBB^T L^T P\right)e\left(t\right) + \varepsilon_2^{-1}\upsilon^T\left(t\right)\upsilon\left(t\right) \tag{25}$$

$$e^T\left(t\right)PD\omega\left(t\right) + \omega^T\left(t\right)D^T Pe\left(t\right) \le \varepsilon_3 e^T\left(t\right)\left(PDD^T P\right)e\left(t\right) + \varepsilon_3^{-1}\omega^T\left(t\right)\omega\left(t\right) \tag{26}$$

Noticing Assumptions 3 and Equation (17), we have:

$$\mu^T\left(t\right)\mu\left(t\right) \le a^2 e^T\left(t\right)e\left(t\right) + a^2\upsilon^T\left(t\right)\upsilon\left(t\right) \tag{27}$$

Substituting Equations (24)–(27) into Equation (23) results in:

$$\dot{V}\left(e\left(t\right)\right) \le e^T\left(t\right)\Pi e\left(t\right) + \left(\varepsilon_1^{-1}a^2 + \varepsilon_2^{-1}\right)\upsilon^T\left(t\right)\upsilon\left(t\right) + \varepsilon_3^{-1}\omega^T\left(t\right)\omega\left(t\right) \tag{28}$$

where:

$$\Pi := (A - LC)^T P + P(A - LC) + P(\varepsilon_1 LL^T + \varepsilon_2 LBB^T L^T + \varepsilon_3 DD^T) P + \varepsilon_1^{-1} a^2 I \tag{29}$$

and it is known from Equation (19) that $\Pi < 0$.

Based on Equation (28), it can be found that:

$$
\begin{aligned}
&\frac{d\left(e^{\beta t} V(e(t))\right)}{dt} \\
&= e^{\beta t}\left[\beta V(e(t)) + \dot{V}(e(t))\right] \\
&\leq e^{\beta t}[\beta \lambda_{\max}(P) e^T(t) e(t) + e^T(t)\Pi e(t) + (\varepsilon_1^{-1} a^2 + \varepsilon_2^{-1}) v^T(t) v(t) + \varepsilon_3^{-1}\omega^T(t)\omega(t)] \\
&= e^{\beta t}[-(\lambda_{\min}(-\Pi) - \beta\lambda_{\max}(P))\|e(t)\|^2 + (\varepsilon_1^{-1} a^2 + \varepsilon_2^{-1})\|v(t)\|^2 + \varepsilon_3^{-1}\|\omega(t)\|^2]
\end{aligned}
\tag{30}
$$

Let $\beta = \dfrac{\lambda_{\min}(-\Pi)}{\lambda_{\max}(P)} > 0$. Integrating both sides from 0 to $T > 0$ and taking the expectation lead to:

$$
\begin{aligned}
e^{\beta T}\mathrm{E}\left[V(e(T))\right] &\leq \lambda_{\max}(P)\sup_{t_0}\mathrm{E}\left[\|e_0\|^2\right] + \left[-(\lambda_{\min}(-\Pi) - \beta\lambda_{\max}(P))\right]\mathrm{E}\left[\int_0^T e^{\beta t}\|e(t)\|^2 \, dt\right] \\
&\quad + \int_0^T\left((\varepsilon_1^{-1} a^2 + \varepsilon_2^{-1})\mathrm{E}\left[\|v(t)\|^2\right] + \varepsilon_3^{-1}\mathrm{E}\left[\|\omega(t)\|^2\right]\right)e^{\beta t} dt \\
&\leq \lambda_{\max}(P)\sup_{t_0}E\left[\|e_0\|^2\right] + \left((\varepsilon_1^{-1} a^2 + \varepsilon_2^{-1})\rho_1 + \varepsilon_3^{-1}\rho_2\right)\int_0^T e^{\beta t} dt \\
&\leq \lambda_{\max}(P)\sup_{t_0}E\left[\|e_0\|^2\right] + \beta^{-1}\rho(e^{\beta T} - 1)
\end{aligned}
\tag{31}
$$

where $\mathrm{E}\left[\|v(t)\|^2\right] \leq \rho_1$, $\mathrm{E}\left[\|\omega(t)\|^2\right] \leq \rho_2$ (according to Assumption 2, $\omega(t)$ and $v(t)$ are bounded random vectors), $\rho = (\varepsilon_1^{-1} a^2 + \varepsilon_2^{-1})\rho_1 + \varepsilon_3^{-1}\rho_2$ and:

$$\mathrm{E}\left[\|e(T)\|^2\right] \leq \lambda_{\min}^{-1}(P)\lambda_{\max}(P)\sup_{t_0}\mathrm{E}\left[\|e_0\|^2\right]e^{-\beta T} + \beta^{-1}\rho\lambda_{\min}^{-1}(P)(e^{\beta T} - 1)e^{-\beta T} \tag{32}$$

Notice that $(e^{\beta T} - 1)e^{-\beta T} < 1$ and let $\alpha := \lambda_{\min}^{-1}(P)\lambda_{\max}(P)$, $\gamma := \beta^{-1}\rho\lambda_{\min}^{-1}(P)$. Since $T > 0$ is arbitrary, the definition of exponential ultimate boundedness in (18) is then satisfied, and this completes the proof of Theorem 2.

5. Simulation Results

In order to show the applicability of the proposed filtering algorithm, we consider the following nonlinear system represented as :

$$
\begin{cases}
\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -0.3 & 0.3 \\ 0 & a(t) \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ b(t) \end{bmatrix}\omega \\
y(t) = x_1(t) + x_1(t)x_2(t) + v^2
\end{cases}
\tag{33}
$$

where $a(t) = -0.3 + 0.03\sin(1 + t)^{-1}$ and $b(t) = 0.2((t + 1)^{-1/2} + 1)$. The random disturbances ω and v obey non-Gaussian, and their distributions are shown in Figure 1. The weights in Equation (14) are

selected as $R_1 = 10$, $R_2 = 2$ and $R_3 = 10$, respectively. The simulation results based on the MEE filter are shown in Figures 2–6, 7a and 8a). And the comparative results between MEE filter and UKF are shown in Figures 7 and 8.

Figure 2 demonstrates that the performance index decreases monotonically with time. In Figures 3 and 4, both the range and PDFs of errors are given, it can be seen that the shapes of PDFs of the tracking errors become narrower and sharper along with the increasing time, which illustrates the dispersions of estimation errors can be reduced. In order to clarify the improvements, the initial and final PDFs are shown in Figures 5 and 6. It can be shown from Figures 3–6 that the proposed improved MEE filter can decrease the uncertainties of the estimation errors and drive the estimation errors approaching to zero. Figure 7 shows that both the MEE filter and UKF make the estimation errors approach to zero. But the estimation errors in Figure7a are more closer to zero with smaller randomness. The variances of the filter gains in two cases are presented in Figure 8. It can be seen that the gains of UKF are not convergent.

From the above analyses, it is obvious that the proposed strategy has better performance than UKF. Therefore, the proposed MEE filter is more suitable for nonlinear stochastic systems with non-Gaussian noises.

Figure 1. Distribution of random noises ω and υ
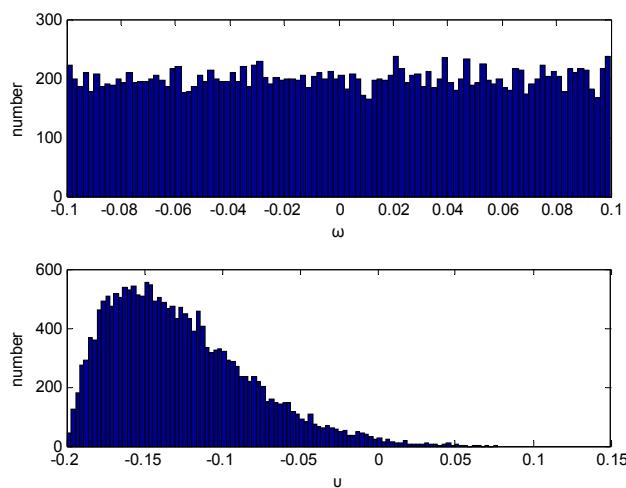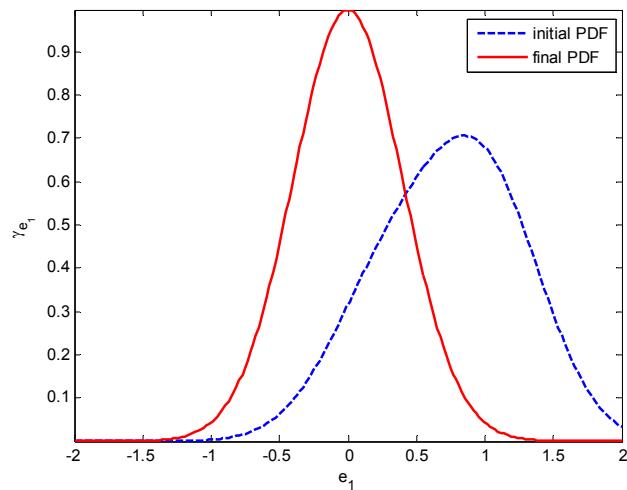

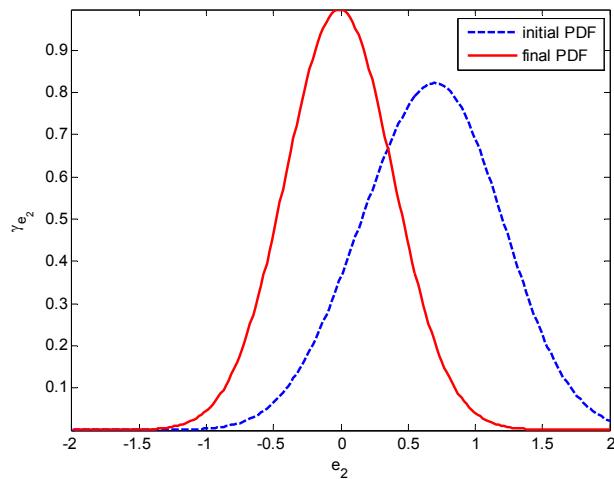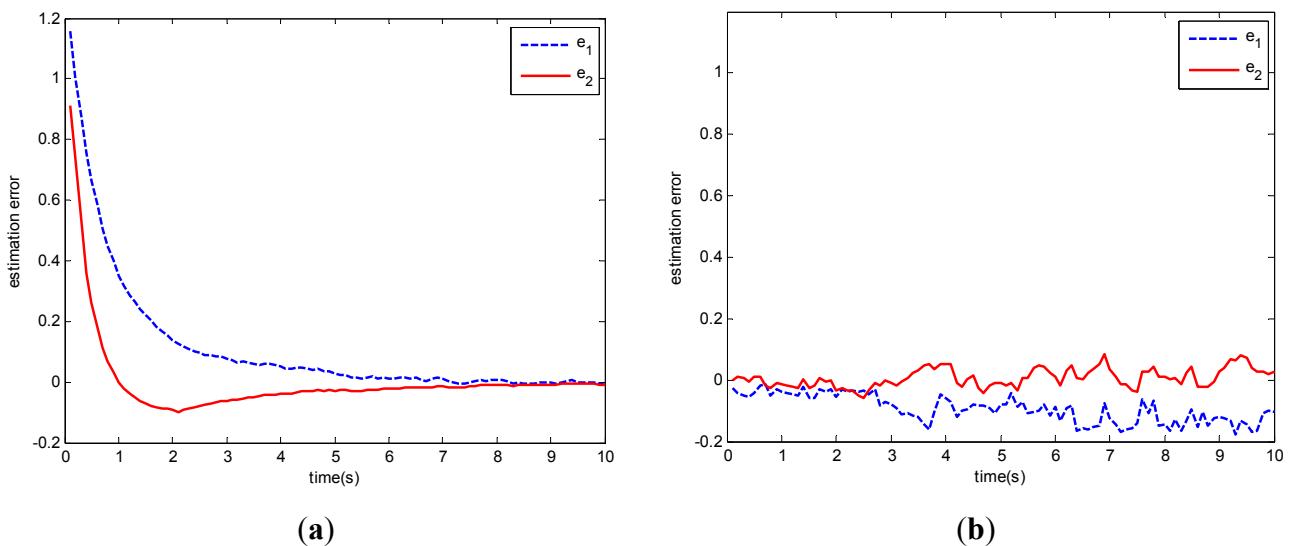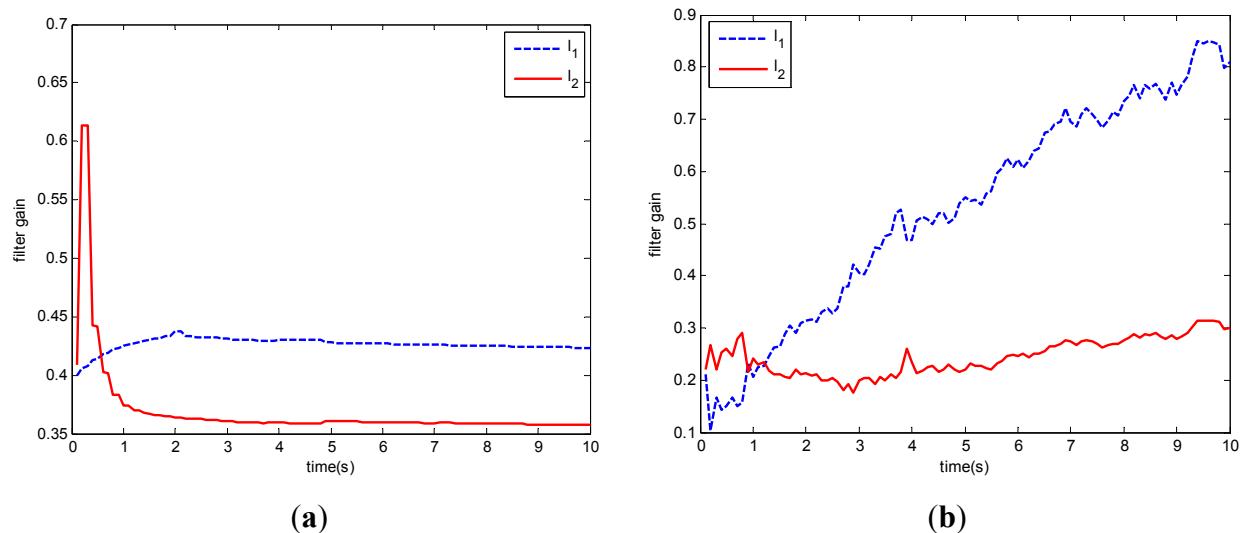**Figure 2.** Performance index.
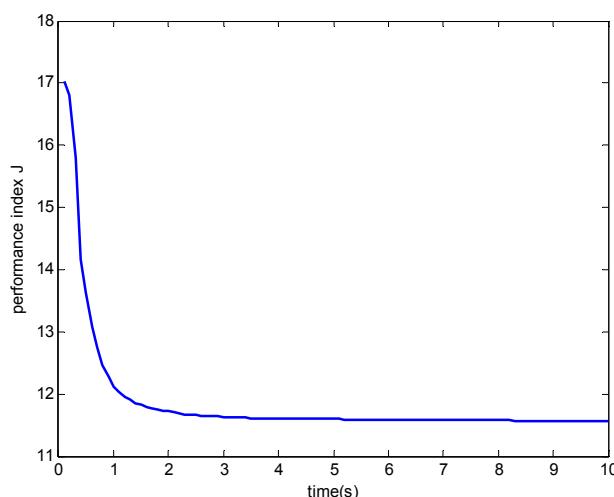
Figure 3. PDF of estimation error e_1
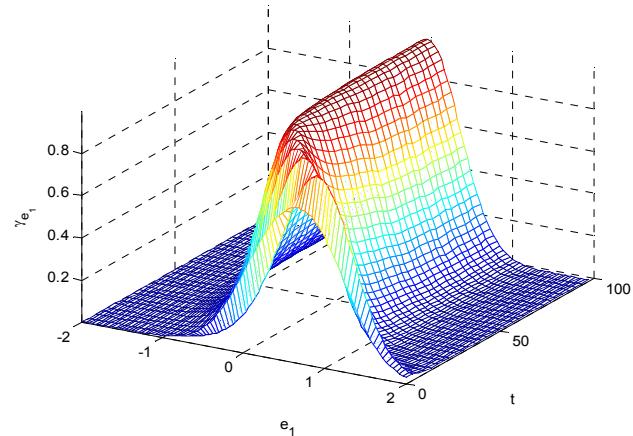
Figure 4. PDF of estimation error e_2.
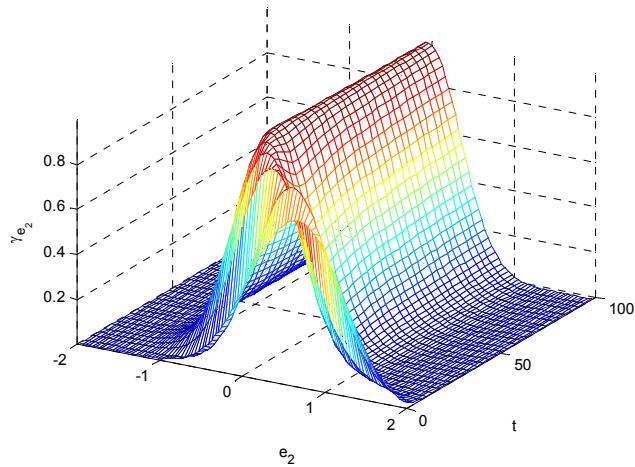
Figure 5. Initial and final PDFs of e_1

Figure 6. Initial and final PDFs of e_2

Figure 7. Estimation errors under the filter: (**a**) MEE filter (**b**) UKF.

(**a**)

(**b**)

Figure 8. Filter gain: (**a**) MEE filter (**b**) UKF.

(**a**)

(**b**)

6. Conclusions

In this paper, based on the improved MEE criterion, optimal filter design problem is studied for multivariate nonlinear continuous stochastic systems with non-Gaussian noises. A novel performance index, including Renyi's entropy of the estimation error, mean value of the squared error and constrains on filter gain, is formulated. By using the generalized density evolution equation, the relationship among the PDFs of estimation error, random disturbances and filter gain matrix is obtained. A recursive optimal filtering algorithm is obtained by minimizing the improved MEE criterion. The exponentially mean-square boundedness condition of the estimation error systems is established. In this work, the proposed MEE filter and UKF are both applied to a continuous nonlinear and non-Gaussian stochastic system. Comparative simulation results demonstrate the superiority of the presented control algorithm.

Conflict of Interest

The authors declare 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. Kalman, R.E.; Bucy, R.S. New results in linear filtering and prediction theory. *Trans. ASME J. Basic Eng.* **1961**, *83*, 95–107.
3. Hoblos, V.G.; Chafouk, H. A gain modification form for Kalman filtering under state inequality constraints. *Int. J. Adapt. Control* **2010**, *24*, 1051–1069.
4. Guo, L.; Wang, H. Fault detection and diagnosis for general stochastic systems using B-spline expansions and nonlinear filters. *IEEE Trans. Circ. Syst. I* **2005**, *52*, 1644–1652.
5. Li, T.; Guo, L. Optimal fault-detection filtering for non-gaussian systems via output PDFs. *IEEE Trans. Syst. Man Cybern. Syst. Hum.* **2009**, *39*, 476–481.
6. Guo, L.; Wang, H. Minimum entropy filtering for multivariate stochastic systems with non-Gaussian noises. *IEEE Trans. Automat. Contr.* **2006**, *51*, 695–700.
7. Zhou, J.; Zhou, D.; Wang, H.; Guo, L.; Chai, T. Distribution function tracking filter design using hybrid characteristic functions. *Automatica* **2010**, *46*, 101–109.
8. Kitagawa, G. Non-Gaussian state-space modeling of nonstationary time-series. *J. Am. Stat. Assoc.* **1987**, *82*, 1032–1063.
9. Kramer, S.C.; Sorenson, H.W. Recursive Bayesian estimation using piece-wise constant approximations. *Automatica* **1988**, *24*, 789–801.

10. Tanizaki, H.; Roberto, S. Mariano, Nonlinear and Non-Gaussian State-space Modelling with Monte-Carlo Simulations. *J. Econometrics* **1998**, *83*, 263–290.

11. Kitagawa, G. Monte carlo filter and smoother for non-Gaussian nonlinear state space models. *J. Comput. Graph. Stat.* **1996**, *5*, 1–25.

12. Cappe, O.; Godsill, S.; Moulines, E. An overview of existing methods and recent advances in sequential Monte Carlo. *Proceedings of the IEEE* **2007**, *95*, 899–924.

13. Vaswani, N. Particle filtering for large dimensional state spaces with multimodal observation likelihoods. *IEEE Trans. Signal Proces.* **2008**, *56*, 4583–4597.

14. Tang, X.; Huang, J.; Zhou, J.; Wei, P. A sigma point-based resampling algorithm in particle filter. *Int. J. Adapt. Control* **2012**, *26*, 1013–1023.

15. Zhao, S.Y.; Liu, F.; Luan, X.L. Risk-sensitive filtering for nonlinear Markov jump systems on the basis of particle approximation. *Int. J. Adapt. Control* **2012**, *26*, 158–170.

16. Erdogmus, D.; Principe, J.C. An error-entropy minimization algorithm for supervised training of nonlinear adaptive systems. *IEEE Trans. Signal Proces.* **2002**, *50*, 1780–1786.

17. Principe, J.C. *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*; John Wiley & Sons: New York, NY, USA, 2010.

18. Chen, B.; Hu, J.; Li, H.; Sun, Z. Adaptive filtering under maximum mutual information criterion. *Neurocomputing* **2008**, *71*, 3680–3684.

19. Li,J.; Chen, J. The principal of preservation of probability and the generalized density evolution equation. *Struct. Saf.* **2008**, *30*, 65–77.

20. Peng, Y.; Li, J. Exceedance probability criterion based stochastic optimal polynomial control of Duffing oscillators. *Int. J. Nonlin. Mech.* **2011**, *46*, 457–469.

21. Li, J.; Peng, Y.; Chen, J. A physical approach to structural stochastic optimal controls. *Probabilist. Eng. Mech.* **2010**, *25*, 127–141.

22. Li, J.; Chen, J. *Stochastic Dynamics of Structures*; John Wiley & Sons: New York, NY, USA, 2009.

23. Wang, Z.; Ho Daniel, W.C. Filtering on nonlinear time-delay stochastic systems. *Automatica* **2003**, *39*, 101–109.

24. Mao, X. Stochastic Differential Equations and Applications; Harwood University: New York, NY, USA, 1997.