

# Generalized Correntropy Criterion-Based Performance Assessment for Non-Gaussian Stochastic Systems

Jinfang Zhang , Guodou Huang  and Li Zhang

School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; 120192227115@ncepu.edu.cn (G.H.); 120192227034@ncepu.edu.cn (L.Z.)

\* Correspondence: jfzhang@ncepu.edu.cn; Tel.: +86-188-1063-3957

**Abstract:** Control loop performance assessment (CPA) is essential in the operation of industrial systems. In this paper, the shortcomings of existing performance assessment methods and indicators are summarized firstly, and a novel evaluation method based on generalized correntropy criterion (GCC) is proposed to evaluate the performance of non-Gaussian stochastic systems. This criterion could characterize the statistical properties of non-Gaussian random variables more fully, so it can be directly used as the assessment index. When the expected output of the given system is unknown, generalized correntropy is used to describe the similarity of two random variables in the joint space neighborhood controlled and take it as the criterion function of the identification algorithms. To estimate the performance benchmark more quickly and accurately, a hybrid-EDA (H-EDA) combined with the idea of “wading across the stream algorithm” is proposed to obtain the system parameters and disturbance noise PDF. Through the simulation of a single loop feedback control system under different noise disturbances, the effectiveness of the improved algorithm and new indexes are verified.



**Citation:** Zhang, J.; Huang, G.; Zhang, L. Generalized Correntropy Criterion-Based Performance Assessment for Non-Gaussian Stochastic Systems. *Entropy* **2021**, *23*, 764. <https://doi.org/10.3390/e23060764>

Academic Editor: Philip Broadbridge

Received: 19 May 2021

Accepted: 15 June 2021

Published: 17 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** performance assessment indicators; hybrid-EDA; non-Gaussian systems; generalized correntropy criterion

## 1. Introduction

With the continuous improvement of automation in the industrial process, the production process is becoming more and more complex. Control loops play the most important role in automation systems. Product quality, operation safety, material consumption and energy consumption are closely related to the performance of control system. Excellent control loop performance will ensure the effectiveness of the control systems, thus ensuring product quality and reducing product cost. The significance of performance assessment is to realize, restore and maintain the optimal performance of the control loops at all times. Therefore, it is of great practical value to evaluate the system performance quickly and accurately in industrial operation. At present, control loop performance assessment (CPA) has become an essential technology to ensure the smooth progress of industrial production [1].

Noise is unavoidable during the performance evaluation [2]. Most of the current performance evaluation methods assume that the noise obeys the Gaussian distribution. The research focuses on the first or second-order moments of the target output (i.e., the mean and variance). Harris [3] proposed a CPA index based on minimum variance control (MVC). This method is regarded as a milestone in the performance assessment research, so we named the index after the researcher, Harris. Later researchers have made great progress on this fundamental. The MVC-based performance evaluation index has been applied to various types of control systems, and this index has been improved for different situations [4–6]. At present, methods of system performance assessment based on MVC are very mature and has a remarkable application effect when the noise disturbance conforms

to Gaussian. However, most of the disturbances in the actual industrial process obey non-Gaussian. The traditional MVC methods are no longer applicable for this case. A more suitable indicator should be chosen to reflect the higher-order statistical characteristics of variables rather than mean and variance in the closed-loop control systems.

For a stochastic system with non-Gaussian noises, the probability density function (PDF) of the output data is approximated by B-spline in the early research direction [7]. A stochastic distributed control algorithm based on minimum entropy control (MEC) was proposed by Zhang and Chu [8]. Information entropy has a more general meaning for any stochastic variable than mean or variance. All higher-order moments (not just second-order) can be minimized by minimizing entropy rather than mean square error. Therefore, the application effect of MEC control strategy is much better than MVC in non-Gaussian systems. On this basis, some significant progress has been made in the field of CPA with non-Gaussian noise [9–14]. When the expected output PDF is not given in advance, it is necessary for the CPA methods of the controlled system based on MEC to identify the associated system model effectively and systematically based on closed-loop data. Minimum entropy criterion was introduced into the feedback control system by Jiang [9], and a new performance assessment method came into being. In work [9], the model of the associated system is identified by the estimation of distribution algorithm (EDA) to obtain the optimal parameters, then the mathematical model and disturbance sequence between output and disturbance is obtained. It is verified that the new benchmark is appropriate for both Gaussian and non-Gaussian systems. However, the method given by Jiang [9] is not detailed enough, the specific calculation methods of minimum entropy for discrete and continuous disturbances are not given. Fortunately, an improved minimum entropy calculation method was proposed by Zhou [10,11] to correct the defects of Jiang's method. In work [10], Zhou and Zhang analyzed the defects of Shannon entropy in CPA process and proposed a new entropy index, rational entropy. The method given in [10] can calculate the theoretical benchmark value. In work [11], the traditional EDA has been improved to identify the system parameters and estimate the noise PDF better and the estimation method of benchmark value is also given. In reference [12], the rational entropy index is used to evaluate a typical cascade control system. However, entropy has the property of translation invariance, that is, the size of entropy is only determined by the shape of its distribution. When the mean of error distribution is not around zero, the wrong assessment result may be caused. For this case, Zhou proposed the performance index of mean-constraint, while a CPA method based on mixture correntropy was proposed by Zhang [13]. The problem of the minimum entropy translation invariance can be solved by adjust the kernel width and weight coefficient of mixture correntropy. Similarly, to avoid the defect of Renyi entropy indicator is insensitive to mean shift, Zhang [14] combined the Renyi entropy benchmark with the mean value to construct a new indicator that can reflect the mean shift of non-Gaussian system. In conclusion, the main breakthrough of the performance assessment method based on MEC is the improvement of the identification algorithm and the selection of some new entropy indicators.

Correntropy is a measure of the generalized similarity of two random variables using information from the statistical properties of higher-order signals [15]. The kernel function of mixture correntropy CPA [13] is Gaussian kernel by default. However, it can only reflect a specific type of noise with Gaussian function, and the shape of entropy cannot be changed freely. Therefore, the Gaussian kernel is not always the best choice [16]. A more general correntropy, generalized correntropy (GC) is proposed. The generalized Gaussian density function is used to replace the traditional Gaussian kernel function for correlation analysis. With greater flexibility, generalized correntropy could describe the statistical characteristics of non-Gaussian random variables better and more fully. In recent years, generalized correntropy has been widely used to design the robust adaptive filters to adapt to different applications [17–20]. For instance, to improve the accuracy and stability of the traditional multi-kernel filtering algorithms, the multi-kernel generalized maximum correntropy (MKGMC) filtering algorithm is developed by Sun [17] under the

generalized maximum correntropy criterion. Wang [18] applied generalized correntropy to the design of sparse Gaussian Hermite orthogonal filters. Qiu [19] replaced the mean square error loss with generalized correntropy loss (GCL) in the original unscented Kalman filter (UKF) framework, which combines the strength of GCL developed in robust information theoretic learning to solve the non-Gaussian interference and the strength of the UKF in handling strong model nonlinearities. Therefore, it is very promising to study the performance assessment of non-Gaussian stochastic systems by replacing entropy criterion with generalized correntropy.

To estimate the performance benchmark better, this paper integrates the idea of wading across the stream algorithm (WSA) into the traditional EDA to improve the local convergence ability of the algorithm. In the iterative process of the algorithm, the center of some excellent individuals is selected to cross with the best solution to fully retain the information of the best individuals. Generalized correntropy is used as the fitness value of H-EDA to identify the parameters of the controlled auto regression and moving average (CARMA) model of the controlled system, which accelerate the algorithm and improves the accuracy. The effectiveness of the hybrid algorithm is verified through the simulation of a single loop feedback control system under different noise disturbances.

In the next section, the CPA index based on MEC is reviewed, and the generalized correntropy index is proposed to overcome the shortcomings of the previous entropy indexes. The third part introduces the performance evaluation process based on the H-EDA in detail. In the fourth section, a detailed simulation of a single input single output (SISO) system is carried out, and the final conclusion will be given in the last part.

## 2. Performance Assessment Based on MEC

### 2.1. Feedback Control Loop

A SISO feedback control system is considered in this paper for convenience of research, the detailed description is given in Figure 1.

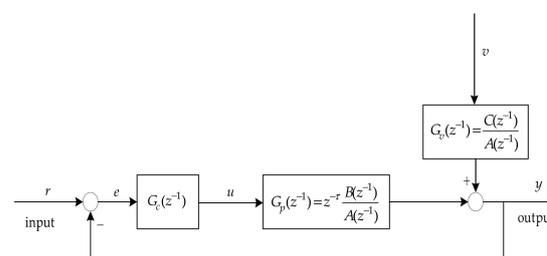


Figure 1. SISO feedback control loop system.

where,  $r$  represents the system set value;  $u$  represents the output of the controller;  $v$  represents stationary independent identically distributed unmeasurable noise.

The system in Figure 1 could be described as a CARMA model as follows,

$$A(z^{-1})y(k) = z^{-\tau}B(z^{-1})u(k) + C(z^{-1})v(k) \tag{1}$$

where  $y(k)$ ,  $u(k)$  and  $v(k)$  are the output, input and unknown noise disturbance of CARMA model process respectively,

$$\begin{cases} A(z^{-1}) = 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{n_a}z^{-n_a} \\ B(z^{-1}) = b_1z^{-1} + b_2z^{-2} + \dots + b_{n_b}z^{-n_b} \\ C(z^{-1}) = 1 + c_1z^{-1} + c_2z^{-2} + \dots + c_{n_c}z^{-n_c} \end{cases} \tag{2}$$

where  $n_a$ ,  $n_b$ ,  $n_c$  and  $\tau$  are the structural parameters of the model;  $n_a$ ,  $n_b$ , and  $n_c$  are the order of  $A(z^{-1})$ ,  $B(z^{-1})$  and  $C(z^{-1})$  respectively,  $\tau$  is the system delay. The parameters  $a_{n_a}$ ,  $b_{n_b}$ , and  $c_{n_c}$  can be estimated by the model identification, provided the structural parameters are obtained.

## 2.2. Minimum Entropy Index

Assuming that the input of the given system is 0, the output is,

$$y_t = \frac{G_v}{1 + G_p G_c} v_t = \frac{G_v}{1 + z^{-\tau} \tilde{G}_p G_c} v_t \quad (3)$$

where,  $\tilde{G}_p$  is the transfer function without delay. By solving the Diophantine equation, the disturbance transfer function  $G_v$  could be further decomposed,

$$G_v(z^{-1}) = F(z^{-1}) + z^{-\tau} R(z^{-1}) = (1 + n_1 z^{-1} + n_2 z^{-2} + \dots + n_{\tau-1} z^{-(\tau-1)}) + z^{-\tau} R(z^{-1}) \quad (4)$$

where,  $F(z^{-1})$  is the impulse response coefficient of the disturbance transfer function  $G_v$  and  $R(z^{-1})$  is the remaining transfer function satisfying the identity (4),

$$y(t) = Fv(t) + Lv(t - \tau) = \underbrace{(n_0 + n_1 z^{-1} + n_2 z^{-2} + \dots + n_{\tau} z^{-(\tau-1)})v(t)}_{\text{feedback-invariant}} + \underbrace{(n_d z^{-\tau} + n_{d+1} z^{-(\tau+1)} + \dots)v(t)}_{\text{feedback-varying}} \quad (5)$$

where,

$$L = \frac{R - F\tilde{G}_p G_c}{1 + z^{-\tau} \tilde{G}_p G_c} \quad (6)$$

The control objective of the minimum entropy controller can be achieved by minimizing the entropy of the output variables when analyzing the linear non-Gaussian system [21,22]. The feedback-invariant only depend on the characteristics of the disturbance in the process rather than the function of the process model or controller. The second term on the right of Equation (5) is the feedback-varying, which means the controller  $G_v$  can determine the entropy value of the process output. Since both sides of Equation (6) are independent, therefore,

$$H(y_t) = H(Fv_t + Lv_{t-\tau}) \geq H(Fv_t) \quad (7)$$

Just like the MVC method, the equal sign in Equation (7) holds only when  $L = 0$ , then the output entropy reaches the minimum, that is the benchmark entropy,

$$H^{\min}(y_t) = H(Fv_t) \quad (8)$$

The CPA based on MEC is to compare the actual system output entropy  $H(y_t)$  with the output entropy  $H^{\min}(y_t)$  under MEC. To sum up, the CPA index based on MEC can be expressed as,

$$\eta = \frac{H^{\min}(y_t)}{H(y_t)} = \frac{H(Fv_t)}{H(y_t)} \quad (9)$$

This indicator is similar to MVC, which is always between 0 and 1. Generally, the closer it is to 1, the closer the system to the ideal case, indicating that the system performs better; conversely, the closer it is to 0, the worse the system performance is, even including unstable control.

## 2.3. Prevenient Entropy Index and Generalized Correntropy

Entropy is a measure to describe the uncertainty of random variables, even for random variables without mean or variance. The size of entropy is only determined by the shape of the distribution rather than its location. There are many ways to describe entropy, such as Shannon entropy (SE), rational entropy (RE), Renyi entropy and delta entropy. In the researches based on the minimum entropy criterion, SE is widely used. For linear Gaussian

systems, *SE* is equivalent to variance. For a continuous random variable  $x$ , Shannon entropy can be expressed as,

$$H_{SE} = - \int \gamma(x) \ln \gamma(x) dx, x \in R \tag{10}$$

where  $\gamma(x)$  is its probability density function (PDF). *SE* is one of the CPA criteria based on MEC system in previous researches [10]. However, it does not meet the “consistency” principle, that is, the results must be the same when entropy could be calculated in many different ways. These characteristics show that *SE* cannot be used as a new evaluation index.

Fortunately, the rational entropy (*RE*) criterion is proposed by Zhou [10], which has most of the properties of Shannon entropy and satisfies the “consistency” principle. The definition of rational entropy is,

$$H_{RE} = - \int \gamma(x) \log \frac{\gamma(x)}{1 + \gamma(x)} dx, x \in R \tag{11}$$

where  $R$  is the domain of random variable  $x$ , and  $\gamma(x)$  is its PDF. *RE* is a suitable benchmark for CPA of linear non-Gaussian systems. By comparing the output rational entropy under MEC process with actual, a CPA index will be obtained.

Of course, *RE* is not perfect. The PDF of variable  $x$  is essential to calculate the output *RE*. If the PDF is known, it is more convenient to use the *RE* criterion to calculate the CPA index. However, when the PDF of the variable is unknown, this method is very complicated in the actual calculation process. For instance, when the system parameters are estimated by the EDA, *RE* as the fitness value, needs to be calculated for several times in each iteration (usually 1000 s), and the operation will be repeated for many times in the identification process. In the actual systems, the parameters and data will be more complex, this method will waste a lot of time. Therefore, this method lacks practical engineering significance.

The mixture correntropy index is adopted by Zhang [13] to solve the problem of translation invariance. However, the Gaussian kernel is applied to the kernel function of mixture correntropy in [13], which is not comprehensive in describing the statistical properties of non-Gaussian random variables [16]. As mentioned before, generalized correntropy is a more appropriate indicator.

Correntropy is a measure of local similarity, which is directly related to the similarity of two random variables in the joint space neighborhood controlled by kernel width. Given two random variables  $X$  and  $Y$ , their correntropy is defined as [16],

$$V_{\sigma}(X, Y) = E[\kappa_{\sigma}(X, Y)] = \int \kappa_{\sigma}(x, y) dF_{XY}(x, y) \tag{12}$$

where  $E[\cdot]$  is the expectation value,  $\kappa_{\sigma}$  is the kernel function with the width of  $\sigma$  ( $\sigma > 0$ ) and  $F_{XY}(x, y)$  denotes the joint distribution function of  $(X, Y)$ . Generally, Gaussian kernel is chosen as the kernel function of correntropy,

$$\kappa_{\sigma}(x, y) = G_{\sigma}(e) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{e^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\lambda e^2) \tag{13}$$

where  $e = x - y$  and  $\lambda = 1/2\sigma^2$  is the kernel parameter. In this paper, a well-known GGD function will be used as the kernel function in correntropy [23],

$$G_{\alpha,\beta}(e) = \frac{1}{2\beta\Gamma(1/\alpha)} \exp\left(-\left|\frac{e}{\beta}\right|^{\alpha}\right) = \gamma_{\alpha,\beta} \exp(-\lambda|e|^{\alpha}) \tag{14}$$

where  $\Gamma(\cdot)$  means the gamma function,  $\alpha > 0$  represents shape parameter;  $\beta > 0$  is bandwidth parameter,  $\lambda = 1/\beta^{\alpha}$  is the kernel parameter and  $\gamma_{\alpha,\beta} = \alpha/(2\beta\Gamma(1/\alpha))$  is the normalization constant.

The GGD function is used as the kernel function in the correntropy,

$$V_{\alpha,\beta}(X, Y) = E[G_{\alpha,\beta}(E_{X-Y})] = E[G_{\alpha,\beta}(X - Y)] \tag{15}$$

where, to distinguish the correntropy with Gaussian kernel, we define it as generalized correntropy (GC). Clearly, it is an extension of correntropy.

If the joint distribution  $f(x, y)$  of  $X$  and  $Y$  is known, the correntropy could be calculated by the following formula,

$$V_{\alpha,\beta}(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} G_{\alpha,\beta}(x - y)f(x, y)dxdy \tag{16}$$

Generally, the joint distribution of  $X$  and  $Y$  is hard to known. Similar to correntropy, generalized correntropy is also calculated by sample estimation,

$$\hat{V}_{N,\alpha,\beta}(X, Y) = \frac{1}{N} \sum_{i=1}^N G_{\alpha,\beta}(x(i) - y(i)) = \frac{1}{N} \sum_{i=1}^N G_{\alpha,\beta}(e(i)) \tag{17}$$

Next, some properties of generalized correntropy are summarized as follows to explain the internal conditions for it to be used as a performance index:

Property 1:  $V_{\alpha, \beta}(X, Y)$  is symmetric, that is  $V_{\alpha, \beta}(X, Y) = V_{\alpha, \beta}(Y, X)$ ;

Property 2:  $V_{\alpha, \beta}(X, Y)$  is positive and bounded:  $0 < V_{N, \alpha, \beta}(X, Y) \leq G_{\alpha, \beta}(0) = \gamma_{\alpha, \beta}$ , which reaches the maximum value if and only if  $X = Y$ ;

Property 3: The generalized correntropy contains the absolute higher moments of error variables,  $E_{X-Y} = X - Y$ ,  $V_{\alpha,\beta}(X, Y) = \gamma_{\alpha,\beta} \sum_{n=0}^{\infty} \frac{(-\lambda)^n}{n!} E[|X - Y|^{\alpha n}]$ .

These characteristics show that generalized correntropy can be applied to CPA. When the expected output of the given system is known, the target PDF is expected to be tracked by the controlled output PDF as much as possible, that is, the expected output  $R_k$  and the actual output  $Y_k$  are as consistent as possible,  $e_k = R_k - Y_k$  is close to 0 infinitely. Ideally, the target PDF is tracked by the output completely, the performance of the control system reaches the optimum, and the generalized correntropy reaches the maximum value  $\gamma_{\alpha, \beta}$ . Therefore, the new CPA index based on generalized correntropy can be defined as,

$$\eta_{GC} = \frac{V_{\alpha\beta}(R_k - Y_k)}{\gamma_{\alpha,\beta}} = \frac{E[G_{\alpha, \beta}(e_k)]}{\gamma_{\alpha,\beta}} \tag{18}$$

where:  $E[G_{\alpha, \beta}(e_k)]$  is the generalized correntropy of  $e_k = R_k - Y_k$ . Obviously,  $0 \leq \eta_{GC} \leq 1$ , closer the index is to 1, the better the control system performance is; otherwise, the system performance is poor and needs to be improved.

When the expected output of the system is unknown, it is essential to identify the associated systems. In this case, generalized correntropy could be combined with the EDA to generate a new identification algorithm.

In data analysis of regression and classification, a measure called the correntropic loss (C-loss) is always used instead of correntropy [24]. The GCL between  $X$  and  $Y$  could be defined as,

$$\begin{aligned} J_{GC-loss}(X, Y) &= G_{\alpha,\beta}(0) - V_{\alpha,\beta}(X, Y) \\ &= \gamma_{\alpha,\beta} - V_{\alpha,\beta}(X, Y) \end{aligned} \tag{19}$$

Obviously, minimizing GCL corresponds to maximizing generalized correntropy.

Assuming that  $\{(x_i, y_i)\}_{i=1}^N$  is a sample taken from  $p_{XY}$ , the estimated value of GCL is,

$$\begin{aligned}\hat{J}_{GC-loss}(X, Y) &= G_{\alpha, \beta}(0) - \hat{V}_{\alpha, \beta}(X, Y) \\ &= \gamma_{\alpha, \beta} - \frac{1}{N} \sum_{i=1}^N G_{\alpha, \beta}(x_i - y_i) \\ &= \gamma_{\alpha, \beta} - \frac{1}{N} \sum_{i=1}^N G_{\alpha, \beta}(e_i)\end{aligned}\quad (20)$$

GCL could be applied as the fitness value of the H-EDA to search for the optimal parameters and estimate the noise PDF of the given system; the specific algorithm will be given in Section 3.

### 3. Improvement and Application of EDA in CPA

In essence, the problem of system parameter identification is an alternative problem of high-dimensional parameter space optimization, so the system parameters can be obtained by the methods of high-dimensional parameters optimization. To obtain the noise PDFs, we have to know the order, time delay and the parameters of the given system.

For the estimation of delay, the simplest and most commonly used method is to analyze the correlation between the input  $u_t$  and output  $y_t$  signals [25], which can be expressed as,

$$\hat{\tau}_d = \max_{\tau} E[y(\tau)u(\tau - \tau_d)] \quad (21)$$

When  $\tau = \tau_d$ ,  $R_{y, u}(\tau)$  reaches the peak, the corresponding time is the estimated value of delay.

The Akaike information criterion [26] is applied to get the order of the system in this paper. For the given CARMA model, AIC criterion is as follows,  $n_c$  is the order of the noise model,

$$AIC(n) = \text{Llg}\sigma_e^2 + a(n_a + n_b + n_c) \quad (22)$$

#### 3.1. EDA and Improved

Estimation of distribution algorithm (EDA) is a randomized search heuristic algorithm, which is used to create a probabilistic model of solution space. This model is updated iteratively based on the quality of the solutions sampled by the model [27]. In recent years, EDA as an optimization method has received great attention, and applied to scheduling, project, machine learning and identification problems successfully [27–30]. Wang [28] proposed a hybrid algorithm named estimation of particle swarm distribution algorithm (EPSDA), which combines PSO (the local search method) with the EDA (the global search method) to improve the efficiency of solving nonlinear bilevel programming problems (NBLP). To overcome the instability of model updating in the iterative process, a novel EDA based on the classic compact genetic algorithm (cGA) is proposed by [27] to optimize the benchmark function ONEMAX. Liang [29] developed a new variant of Gaussian EDA (GEDA) to solve the problem of premature convergence in traditional GEDA. In the process of CPA, EDA is mainly devoted to estimate the parameters of the given systems. Through appropriate improvements, the given system can be identified faster and more accurately, so as to obtain a more accurate benchmark.

#### 3.2. Parameter Identification Based on Hybrid-EDA

As the previous work shows, the traditional EDA still has some shortcomings in the optimization process, such as a large amount of calculation, poor accuracy, etc. Therefore, based on the traditional EDA, the space for parameter identification is first obtained through preliminary estimation. To improve the local convergence ability, incorporating WSA [31] ideas into the EDA. In the iterative process, the centers of some excellent individuals are selected to cross-operate with the best individuals, the information of outstanding individuals could be retained more fully.

### 3.2.1. Acquisition of Parameter Identification Space

The initialization of parameter space is of positive significance to improve the identification accuracy and shorten the time required, so it is necessary to estimate the model parameters preliminarily. The recursive extended least squares (RELS) algorithm is selected in this paper.

The above CARMA model can be expressed as,

$$y(k) = \varphi^T(k)\theta + v(k) \tag{23}$$

where,

$$\begin{cases} \varphi(k) = [-y(k-1), \dots, -y(k-n_a), u(k-\tau), \dots, u(k-\tau-n_b), v(k-1), \dots, v(k-n_c)]^T \\ \theta = [a_1, \dots, a_{n_a}, b_0, \dots, b_{n_b}, c_1, \dots, c_{n_c}]^T \in R^{(n_a+n_b+1+n_c) \times 1} \end{cases} \tag{24}$$

Since  $v(k)$  is unmeasurable, it can only be replaced by its estimated value  $\hat{v}(k)$ ,

$$\hat{v}(k) = y(k) - \hat{y}(k) = y(k) - \hat{\varphi}^T(k)\hat{\theta} \tag{25}$$

where,

$$\begin{cases} \hat{\varphi}(k) = [-y(k-1), \dots, -y(k-n_a), u(k-\tau), \dots, u(k-\tau-n_b), \hat{v}(k-1), \dots, \hat{v}(k-n_c)]^T \\ \hat{\theta} = [\hat{a}_1, \dots, \hat{a}_{n_a}, \hat{b}_0, \dots, \hat{b}_{n_b}, \hat{c}_1, \dots, \hat{c}_{n_c}]^T \in R^{(n_a+n_b+1+n_c) \times 1} \end{cases} \tag{26}$$

The purpose is to obtain the parameter vector  $\hat{\theta}$  when the objective function  $J(\hat{\theta})$  reaches the minimum value.  $J(\hat{\theta})$  is defined as follows,

$$J(\hat{\theta}) = \sum_{k=1}^L v^2(k) = \sum_{k=1}^L [y(k) - \varphi^T(k)\hat{\theta}]^2 \tag{27}$$

Based on the estimation of Gaussian system parameters, the initial estimates of variance  $\hat{\sigma}_v$  and parameters  $\hat{\theta}$  could be obtained, then the parameter identification space  $\Omega_W$  could be taken as  $\hat{\theta} \pm 3\hat{\sigma}_v$ .

### 3.2.2. The Idea of Wading Across Stream Algorithm

Wading across stream algorithm (WSA) [31] is inspired by the idea of “crossing the river with stones”. This algorithm examines the “shore” carefully to select an initial starting point firstly, several solutions are randomly searched out in the neighborhood near the starting point to get the best one as the iterative result. Then take this result as the starting point and continue search for several solutions in the nearby neighborhood, select the optimal solution as the third iterations result, and so on until the termination condition is satisfied. The WSA is similar to the simulated annealing algorithm but the effect is better, and the algorithm idea is simple with good local convergence ability, so we integrate this algorithm idea into the EDA to improve its computational efficiency.

- (1) Choose the initial starting point carefully

When the parameter identification space is known, the initially selected solution will have a greater impact on the entire algorithm. Firstly, generate  $R$  solutions in the parameter space with uniform random values, and select the optimal solution as the initial. For the parameter vector  $\hat{\theta} = [\hat{a}_1, \dots, \hat{a}_{n_a}, \hat{b}_0, \dots, \hat{b}_{n_b}, \hat{c}_1, \dots, \hat{c}_{n_c}]$  to be identified, it is necessary to uniformly and randomly take values in the interval  $[\hat{\theta} - 3\hat{\sigma}_v, \hat{\theta} + 3\hat{\sigma}_v]$  to generate  $R$  solutions (vectors) to form a parameter space  $A$ ,

$$A = [\hat{\theta}_1; \hat{\theta}_2; \dots; \hat{\theta}_R] = \begin{bmatrix} \hat{a}_{11} & \cdots & \hat{a}_{1n_a} & \hat{b}_{10} & \cdots & \hat{b}_{1n_b} & \hat{c}_{11} & \cdots & \hat{c}_{1n_c} \\ \hat{a}_{21} & \cdots & \hat{a}_{2n_a} & \hat{b}_{20} & \cdots & \hat{b}_{2n_b} & \hat{c}_{21} & \cdots & \hat{c}_{2n_c} \\ \vdots & \ddots & \vdots \\ \hat{a}_{R1} & \cdots & \hat{a}_{Rn_a} & \hat{b}_{R0} & \cdots & \hat{b}_{Rn_b} & \hat{c}_{R1} & \cdots & \hat{c}_{Rn_c} \end{bmatrix} \tag{28}$$

To calculate the fitness value of these  $R$  solutions, select the optimal individual  $\hat{\theta}^*$  as the initial starting point according to the fitness value. For ease of description, denote the starting point is as  $\hat{\theta}^* = [x_1^*, x_2^*, \dots, x_n^*]$ , where  $n = na + nb + nc + 1$ .

(2) Search strategy

Due to the complexity of parameter identification in high-dimensional space, the design of an effective local search strategy could improve the quality of the solution significantly. According to the idea of “crossing the river with stones”, when you touch a “stone”, you must take that “stone” as the starting point to explore other “stones” around it. Similarly, search for  $m$  individuals in the neighborhood radius  $L_k$  around the starting point  $\hat{\theta}^*$  to form a new parameters space  $B$ ,

$$B = [\hat{\theta}'_1; \hat{\theta}'_2; \dots; \hat{\theta}'_m] = \begin{bmatrix} \hat{a}'_{11} & \dots & \hat{a}'_{1na} & \hat{b}'_{10} & \dots & \hat{b}'_{1nb} & \hat{c}'_{11} & \dots & \hat{c}'_{1nc} \\ \hat{a}'_{21} & \dots & \hat{a}'_{2na} & \hat{b}'_{20} & \dots & \hat{b}'_{2nb} & \hat{c}'_{21} & \dots & \hat{c}'_{2nc} \\ \vdots & \ddots & \vdots \\ \hat{a}'_{m1} & \dots & \hat{a}'_{mna} & \hat{b}'_{m0} & \dots & \hat{b}'_{mnb} & \hat{c}'_{m1} & \dots & \hat{c}'_{mnc} \end{bmatrix} = \begin{bmatrix} x'_{11} & x'_{12} & \dots & x'_{1n} \\ x'_{21} & x'_{22} & \dots & x'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x'_{m1} & x'_{m2} & \dots & x'_{mn} \end{bmatrix} \tag{29}$$

where  $x'_{ij} = x_j^* + L_k \cdot r_{ij}$ ,  $r_{ij}$  represents uniformly distributed random numbers in the interval  $[-1, 1]$ . Generally, the initial value  $L_0$  of neighborhood radius is taken as one tenth of the whole parameters space, that is  $L_0 = (0.1 \sim 0.3)\hat{\sigma}_v$ . To accelerate the later convergence of the algorithm, reduce the neighborhood radius gradually with the iterative process. The simplest method is to set  $L_k = \alpha L_{k-1}$ ,  $\alpha$  is the shrinkage coefficient, generally  $[0.90, 0.99]$ . The latest optimal individual  $\hat{\theta}^*$  can be obtained by calculating the fitness values of these  $m$  solutions.

3.2.3. Crossover Operation

The optimal individual obtained by traditional identification algorithm only contains the information of one individual, which ignores the information contained in other excellent individuals to a certain extent. The improved method is crossover operation, sort the above  $m$  individuals according to the fitness value, select the top  $N^*$  ( $N^* < m$ ) excellent solutions (individuals with the best fitness value)  $D$ ,

$$D = [\hat{\theta}''_1; \hat{\theta}''_2; \dots; \hat{\theta}''_{N^*}] = \begin{bmatrix} \hat{a}''_{11} & \dots & \hat{a}''_{1na} & \hat{b}''_{10} & \dots & \hat{b}''_{1nb} & \hat{c}''_{11} & \dots & \hat{c}''_{1nc} \\ \hat{a}''_{21} & \dots & \hat{a}''_{2na} & \hat{b}''_{20} & \dots & \hat{b}''_{2nb} & \hat{c}''_{21} & \dots & \hat{c}''_{2nc} \\ \vdots & \ddots & \vdots \\ \hat{a}''_{N^*1} & \dots & \hat{a}''_{N^*na} & \hat{b}''_{N^*0} & \dots & \hat{b}''_{N^*nb} & \hat{c}''_{N^*1} & \dots & \hat{c}''_{N^*nc} \end{bmatrix} = \begin{bmatrix} x''_{11} & x''_{12} & \dots & x''_{1n} \\ x''_{21} & x''_{22} & \dots & x''_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x''_{N^*1} & x''_{N^*2} & \dots & x''_{N^*n} \end{bmatrix} \tag{30}$$

Calculate the center point of  $D$ ,

$$\begin{cases} \bar{\theta} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n]^T \\ \bar{x}_i = \frac{1}{N^*} \cdot \sum_{j=1}^{N^*} x_{ji} \end{cases} \tag{31}$$

Then, the center point is crossed with the best solution, and the optimal individual  $\hat{\theta}^*$  is modified as follows,

$$\hat{\theta}^* = a\hat{\theta}^* + (1 - a)\bar{\theta} \tag{32}$$

where:  $a$  is a random number between 0 to 1.

Through the crossover operation, the information of excellent individuals could be retained to the maximum extent, which can avoid falling into the local optimum.

### 3.2.4. Selection of Fitness Value

From Equation (23), for the CARMA model in this paper, the error at time  $k$  could be defined as,

$$e(k) = y(k) - \hat{y}(k) = y(k) - \hat{\varphi}^T(k)\hat{\theta} \tag{33}$$

For the estimated error sequence  $e = [e_1, e_2, \dots, e_L]$ , the GCL could be defined as,

$$\hat{J}_{GC-loss}(e) = \gamma_{\alpha, \beta} - \frac{1}{N} \sum_{i=1}^N G_{\alpha, \beta}(e) \tag{34}$$

The optimal model parameters and the relationship between output and disturbance could be obtained by minimizing GCL. When the GCL reaches the minimum, the corresponding parameters are also optimal, namely,

$$\hat{\theta}^{opt} = \arg \min_{\theta \in \Omega_W} J_{GC}(e_k) \tag{35}$$

where:  $\Omega_W$  is the parameter identification space.

### 3.2.5. Algorithm Summary

The following Algorithm 1 is the program steps of the H-EDA identification algorithm.

---

**Algorithm 1** Program steps for the identification algorithm.

---

1. Describe the system by CARMA model; estimate the system delay  $\tau$  by analyzing the correlation between  $u(t)$  and  $y(t)$ ; determine the model order  $(n_a, n_b, n_c)$ .
  2. Preliminary estimate. Rough model parameters are obtained by RELS, determine the identification space  $\Omega_W$  as  $\hat{\theta} \pm 3\hat{\sigma}_v$ .
  3. Choose the initial starting point carefully. Randomly generate R individuals from the parameter space  $\Omega_W$ ,  $A = [\hat{\theta}_1; \hat{\theta}_2; \dots; \hat{\theta}_R]$ , calculate the fitness value (error entropy) of these solutions in A and select the optimal one  $\hat{\theta}_0^*$  according to the fitness value (corresponding error entropy is minimum).
  4. Search for  $m$  individuals  $B_0 = [\hat{\theta}_1; \hat{\theta}_2; \dots; \hat{\theta}_m]$  within the neighborhood radius  $L_0$  of  $\hat{\theta}_0^*$ ;  $I = 1$ ;  
**While**  $I \leq n_{max}$ 
    - (1) Calculate the fitness value (error entropy) of  $m$  individuals and sorted  $C_I = [\hat{\theta}_1^{(I)*}; \hat{\theta}_2^{(I)*}; \dots; \hat{\theta}_m^{(I)*}]$  based on the fitness value. Extract the optimal one  $\hat{\theta}_I^*$  and top  $N^*$  individuals  $D_I = [\hat{\theta}_1^{(I)*}; \hat{\theta}_2^{(I)*}; \dots; \hat{\theta}_{N^*}^{(I)*}]$  from  $C_I$ ;
    - (2) Modify the optimal individual  $\hat{\theta}_I^*$ . Calculate the center point (mean value)  $\bar{\theta} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n]^T$  of  $D_I$ . Modify the optimal individual as,  $\hat{\theta}_I^* = a\hat{\theta}_I^* + (1 - a)\bar{\theta}$ ,  $a = rand$
    - (3) **If** the termination conditions is met  
 The difference between two adjacent iterations error entropy is less than 0.0001, end the cycle.  
**Else**  
 $L_I = \alpha L_{I-1}$ ;  
 Search for  $m$  individuals  $B_I = [\hat{\theta}_1^{(I)}; \hat{\theta}_2^{(I)}; \dots; \hat{\theta}_m^{(I)}]$  within the neighborhood radius  $L_I$  of  $\hat{\theta}_I^*$ ;  
 $I = I + 1$ ;  
**End**
- End**

5. Get the estimation of parameters  $\hat{\theta}$  and noise PDF  $\hat{v}$ .
- 

After the coefficient  $\hat{F}$  and noise  $\hat{v}_t$  of the feedback-invariant are obtained, the benchmark entropy and output entropy could be obtained by Equations (8) and (11), and the CPA index could be obtained by Equation (9),

$$\hat{\eta} = \frac{H(y_t^{mjc})}{H(y_t)} = \frac{H^{\min}(\hat{F}\hat{v}_t)}{H(y_t)} \tag{36}$$

### 3.3. Algorithm Validation and Sensitivity Analysis of Initial Parameters

In this section, a model example is given to illustrate the initial parameters of improved algorithm, which further analyses the H-EDA and traditional EDA, consider the following system,

$$y(k) = \frac{1 + 1.5z^{-1} + 0.9z^{-2}}{1 - 1.7z^{-1} + 0.7z^{-2}}v(k) \tag{37}$$

For the above example, the parameter vector to be identified is  $\theta = [-1.7, 0.7, 1.5, 0.9]$ . Most of the initial parameters are based on the previous works [12–14,31–33]. Generally, in the EDA, the number of individuals in the initial population is  $N = 1000$ ; the optimal number of individuals for each iteration is  $m = 200$ ; and the maximum number of iterations is  $n_{max} = 120$ . In WSA, the number of excellent individuals for crossover operation is  $N^* = 30, R = 800$ .

In this section, we focus on the sensitivity of shrinkage coefficient  $\alpha$  and the initial value of neighborhood radius  $L_0$ . As mentioned before, the initial neighborhood radius is generally  $L_0 = (0.1 \sim 0.3)\hat{\sigma}_v$ , and the shrinkage coefficient  $\alpha$  is between 0.90 to 0.99. To find a proper set of parameters as much as possible, we have carried out a large number of experiments and some representative results are selected and attached in Table 1. In addition, we analyze the significance of  $a$  to the algorithm,  $a$  is a random number between  $[0, 1]$ . If  $a == 1$ , it means that the crossover operation is not carried out. A novel parameter vector is defined as  $P = [L_0, \alpha, a]$ .

**Table 1.** Model identification results based on the parameters sensitivity analysis (partial).

Project	Parameter Configuration		Identification Results	Iterations		
EDA	$N = 1000, m = 200$		$[-1.4922, 0.5527, 1.6509, 0.6267]$	120		
H-EDA	$L_0 = 0.1\hat{\sigma}_v$	$\alpha = 0.93$	$a = rand$ $a == 1$	$[-1.7209, 0.7241, 1.4680, 0.8668]$ $[-1.7398, 0.7529, 1.4099, 0.7911]$	75 105	
			$\alpha = 0.95$	$a = rand$ $a == 1$	$[-1.7309, 0.7227, 1.4219, 0.8113]$ $[-1.6675, 0.6728, 1.4003, 0.7849]$	82 109
		$\alpha = 0.97$		$a = rand$ $a == 1$	$[-1.7273, 0.7305, 1.4085, 0.8173]$ $[-1.6631, 0.6645, 1.4204, 0.8054]$	91 120
			$L_0 = 0.15\hat{\sigma}_v$	$\alpha = 0.93$	$a = rand$ $a == 1$	$[-1.7197, 0.7218, 1.5160, 0.8851]$ $[-1.7749, 0.7735, 1.4595, 0.8299]$
		$\alpha = 0.95$			$a = rand$ $a == 1$	$[-1.6919, 0.6975, 1.5009, 0.9104]$ $[-1.7181, 0.7184, 1.5090, 0.8544]$
				$\alpha = 0.97$	$a = rand$ $a == 1$	$[-1.7304, 0.7310, 1.5164, 0.8964]$ $[-1.6643, 0.6654, 1.4698, 0.8605]$
	$L_0 = 0.2\hat{\sigma}_v$	$\alpha = 0.93$			$a = rand$ $a == 1$	$[-1.7112, 0.7176, 1.3935, 0.8291]$ $[-1.7672, 0.7689, 1.1863, 0.5712]$
				$\alpha = 0.95$	$a = rand$ $a == 1$	$[-1.6734, 0.6737, 1.5061, 0.9069]$ $[-1.7462, 0.7504, 1.4771, 0.8579]$
		$\alpha = 0.97$			$a = rand$ $a == 1$	$[-1.6823, 0.7446, 1.4500, 0.8403]$ $[-1.7769, 0.7789, 1.4058, 0.8304]$

The test results show that the H-EDA is obviously superior to the traditional EDA. The global search capability of the algorithm is mainly determined by the neighborhood radius  $L_k$ . Too small a value of  $L_k$  prompts the algorithm to get trapped in local optima, while too large of a value of  $L_k$  slows down the algorithm;  $\alpha$  can accelerate the later convergence of the algorithm, but too small of an  $\alpha$  leads to rapid convergence and bad accuracy; the main function of  $a$  is to reduce iterations and prevent local optimality. In short, the setting of initial parameters must rely on a large number of experiments and the summary of

previous works. For this test case, the optimal initial parameter vector can be selected as  $P = [0.15\hat{\sigma}_v, 0.95, rand]$ .

#### 4. Simulation and Verification

To verify the effectiveness of the identification algorithm and the rationality of the performance assessment index, the following system is considered and the numerical simulation is carried out with Gaussian and non-Gaussian noise signals.

$$y(t) = u(t-2) + \frac{1-0.2z^{-1}}{1-z^{-1}}v(t) \quad (38)$$

The transfer function of the controller is as follows,

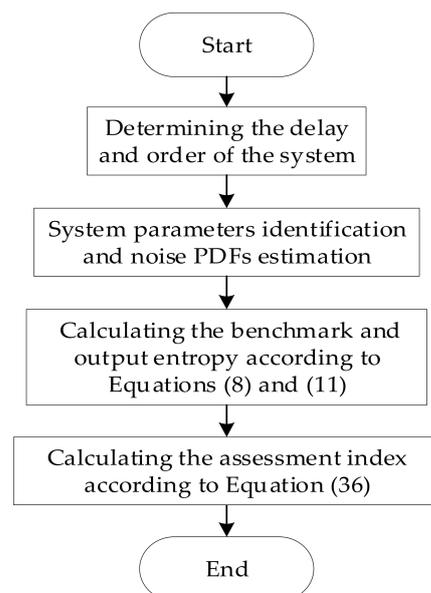
$$G_c = \frac{K}{1-0.2z^{-1}-0.8z^{-2}} \quad (39)$$

It is easy to know that the parameter vector of the system is  $\theta = [-1, 1, -1, -0.2]$ , system delay  $\tau = 2$ , feedback-invariants  $F = [1, 0.8]$ , and the controller gain in the simulation is  $K = 1.2$ .

For more precise description, the follow-up content is divided into the following parts: in Section 4.1, the effect of the improved algorithm and the improved benchmark entropy on the performance evaluation are explained respectively when the expected output is unknown, and their effectiveness are verified respectively. In Section 4.2, the tracking index is used to evaluate the system performance when the expected output of the system is known.

##### 4.1. Simulation When Expected Output Distribution Is Unknown

It can be concluded from Section 3 that the system parameter identification and noise PDF estimation should be carried out firstly when the expected output is unknown, then the benchmark value could be obtained according to the feedback invariant coefficient estimation  $\hat{F}$  and noise estimation  $\hat{\sigma}$ , and the performance evaluation index can be obtained finally. The specific process is shown in Figure 2.



**Figure 2.** Flow chart of performance assessment.

According to the analysis ideas in Section 3.3, the parameters of the algorithms in this section are set as follows:

In the EDA, the number of individuals in the initial population is  $N = 1000$ , the optimal number of individuals for each iteration is  $m = 200$ , and the maximum number of iterations is  $n_{max} = 120$ . Correspondingly, the parameters of H-EDA are set as,  $N^* = 30$ ,  $L_0 = 0.15\hat{\sigma}_v$ ,  $\alpha = 0.95$ ,  $R = 800$ , the rests are consistent with the EDA. As the comparison algorithm, the parameters of the improved PSO (IPSO), particles  $N = 50$ , acceleration factor  $c_1 = 2$ ,  $c_2 = 2$ ,  $V_{max} = 0.5$ .

The termination condition is  $|J^{(l)} - J^{(l-1)}| \leq 0.0001$  or reaches the specified number of iterations  $n_{max}$ ,  $J^{(l)}$  is the  $l$ -generation GCL.

The noises distribution are subject to normal  $N(0, 0.255)$ , Beta  $B(2, 9)$ , exponential  $E(0.5)$  and bimodal,

$$v \sim f(x) = a \cdot \frac{r}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + b \cdot \frac{1-r}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} \tag{40}$$

where  $\mu_1 = -3$ ,  $\mu_2 = 3$ ,  $\sigma_1 = 1$ ,  $\sigma_2 = 0.4$ ,  $r = 0.7$ .

- (1) When the fitness value is fixed as the GCL, compare the H-EDA with the traditional EDA and the IPSO.

Combine the above three algorithms with GCL, respectively. The probability distributions of the four kinds of noise are shown in Figure 3. It can be determined that the other three kinds of noise are non-Gaussian except for normal distribution. The results of the system parameter identification are shown in Table 2.

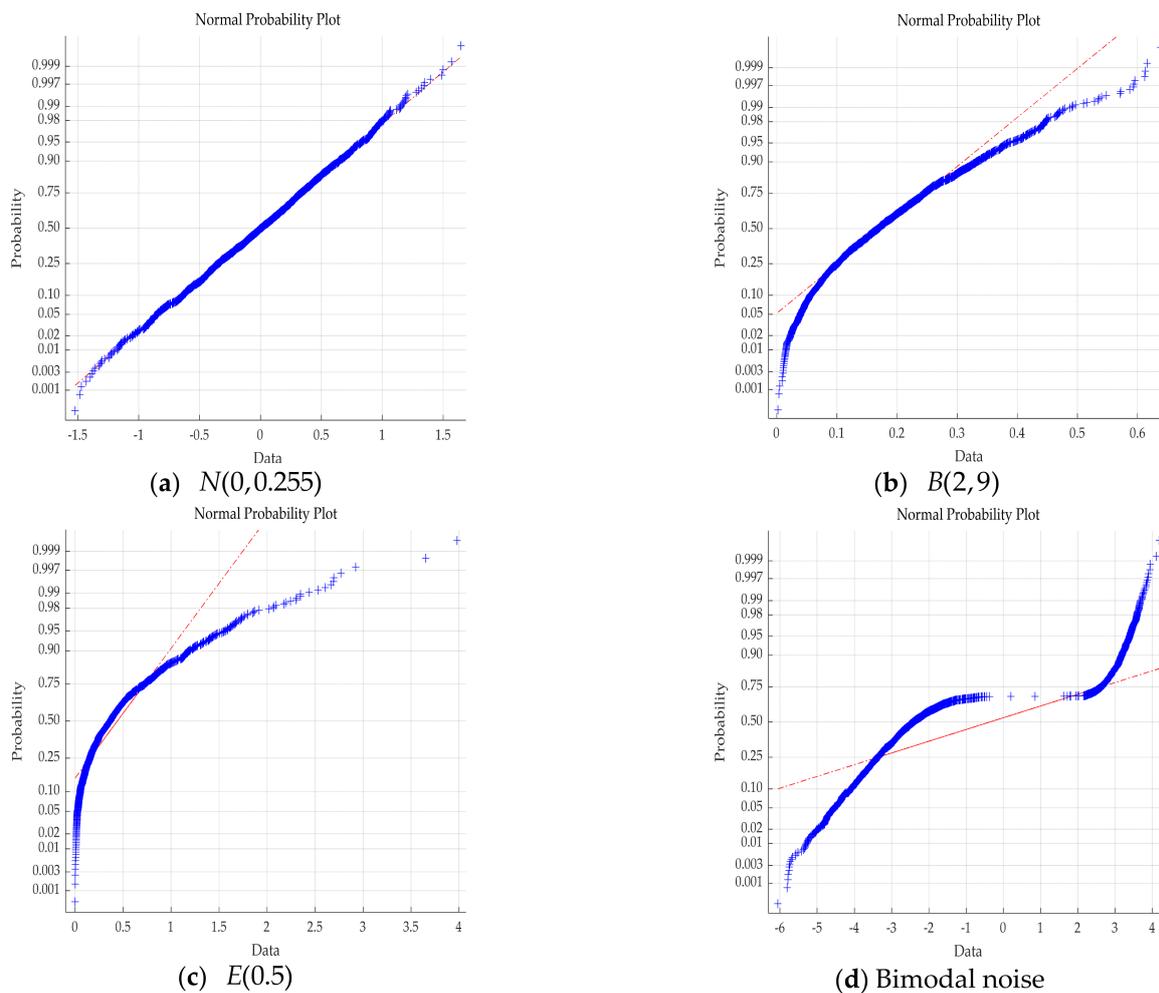


Figure 3. Probability distribution of different noises.

The results in Table 2 show that the H-EDA could estimate the parameters of CARMA model under different noises when the fitness value is determined, and the identification accuracy and speed are better. Moreover, the fitness value of the algorithm is the generalized correntropy loss function, which proves the effectiveness of the novel criterion in the identification process. It is evident that the noises distribution estimated by the H-EDA in Figures 4 and 5 are very close to the real values. IPSO, as the comparison, is inferior to the H-EDA and cannot estimate the noise of bimodal distribution. Therefore, the superiority of the H-EDA is proved by the simulation results.

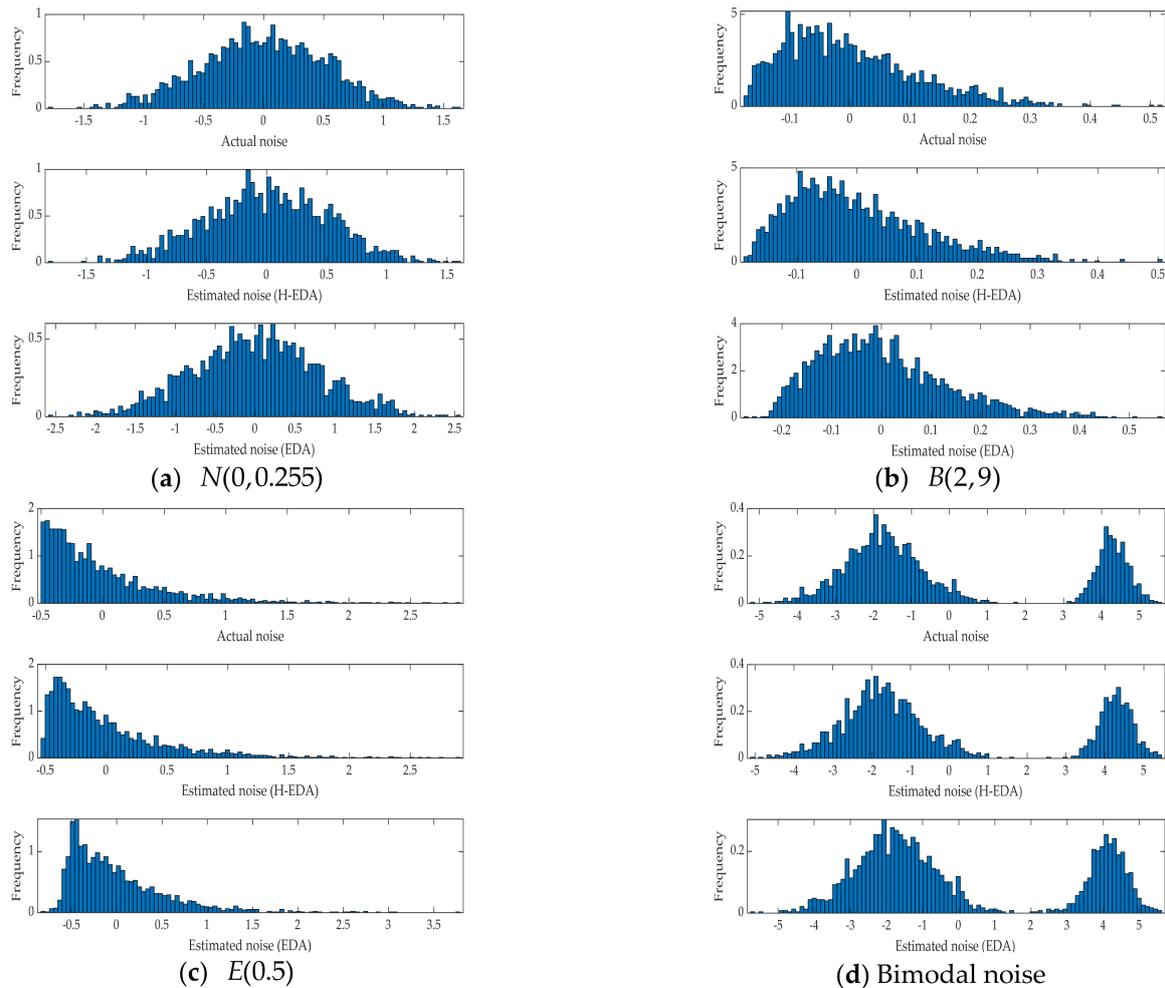
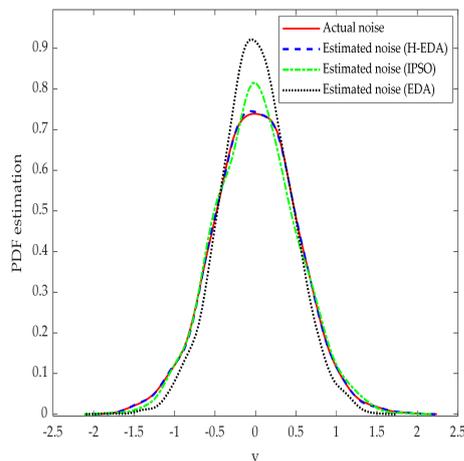


Figure 4. Actual and estimated probability distribution histograms of different noises.

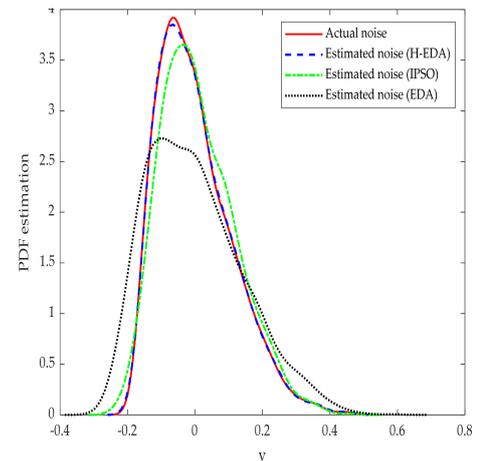
According to the noise PDF  $\hat{v}_t$  and the feedback invariant coefficient  $\hat{F}_t$ , we can get the performance assessment indexes, as shown in Table 3, where  $\eta_{ME}$  represents the theoretical benchmark index;  $\hat{\eta}_{MJ-EDA}$  is the index obtained by identification results of the traditional EDA based on GCL;  $\hat{\eta}_{MJ-H-EDA}$  corresponds to the H-EDA index. Obviously,  $\hat{\eta}_{MJ-H-EDA}$  is closer to the theoretical value, which proves the effectiveness of the hybrid-algorithm again.

**Table 2.** Identification results of two algorithms under different noise distributions.

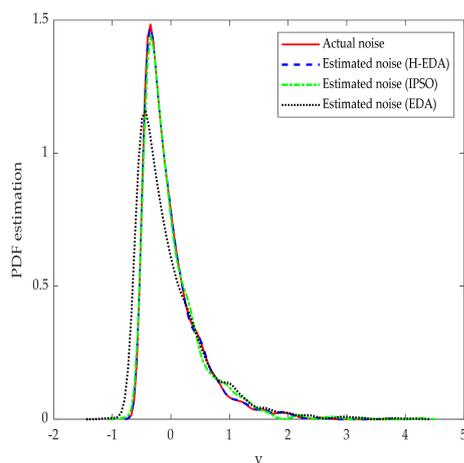
	Actual Value	<i>N</i> (0, 0.255)		<i>B</i> (2, 9)	
		EDA	H-EDA	EDA	H-EDA
$a_1$	-1	-0.9342	-0.9873	-1.0529	-1.0027
$b_0$	1	0.9186	0.9968	1.0848	1.0003
$b_1$	-1	-0.9337	-0.9816	-0.9466	-0.9992
$c_1$	-0.2	-0.0879	-0.1991	-0.2365	-0.2014
$F$	[1, 0.8]	[1, 0.7011]	[1, 0.7782]	[1, 0.8596]	[1, 0.8035]
Time(s)		29.9517	15.2724	35.4492	13.6584
		<i>E</i> (0.5)		Bimodal Noise	
		EDA	H-EDA	EDA	H-EDA
$a_1$	-1	-0.9158	-0.9884	-0.9173	-1.0097
$b_0$	1	0.9389	0.9961	0.9377	1.0643
$b_1$	-1	-0.9323	-0.9924	-0.8946	-1.0853
$c_1$	-0.2	-0.0865	-0.1935	-0.1556	-0.2037
$F$	[1, 0.8]	[1, 0.7595]	[1, 0.7857]	[1, 0.6987]	[1, 0.8138]
Time(s)		30.9885	14.2616	29.6255	14.7962



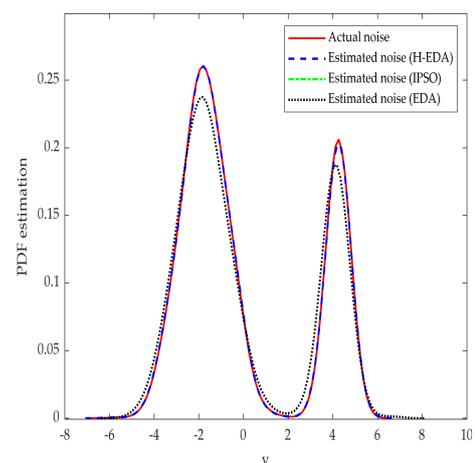
(a)  $N(0, 0.255)$



(b)  $B(2, 9)$



(c)  $E(0.5)$



(d) Bimodal noise (the IPSO is invalid)

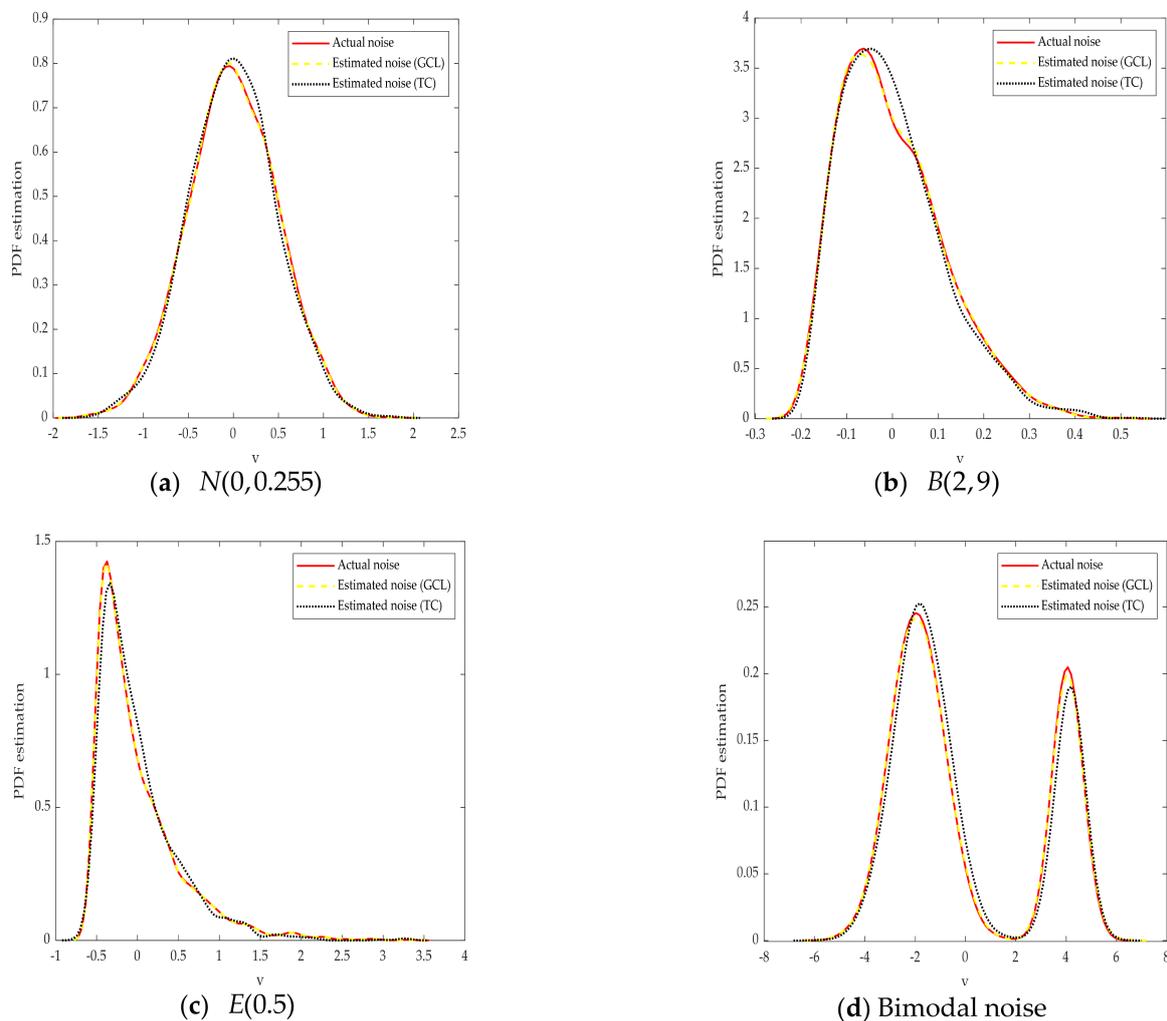
**Figure 5.** Actual and estimated PDF of different noises.

**Table 3.** CPA indexes of two comparison algorithms.

Index	$N(0, 0.255)$	$B(2, 9)$	$E(0.5)$	Bimodal Noise
$\eta_{ME}$	0.8809	0.9357	0.9408	0.9332
$\hat{\eta}_{MJ-EDA}$	0.8345	0.9016	0.9134	0.9267
$\hat{\eta}_{MJ-H-EDA}$	0.8794	0.9388	0.9399	0.9336

- (2) When the identification algorithm is fixed as the H-EDA, compare the traditional correntropy (TC) with the generalized correntropy criterion (GCL).

In the previous section, we have proved the effectiveness of generalized correntropy criterion in the performance assessment process. Next, we compare the superiority of GCL with TC based on the H-EDA. The parameters of algorithms, identified objects and noise PDFs are consistent with (1). The simulation results are illustrated in Table 4 and Figure 6.



**Figure 6.** Actual and estimated PDF with different noises.

**Table 4.** Estimated parameters and evaluation indexes under two different entropy benchmarks.

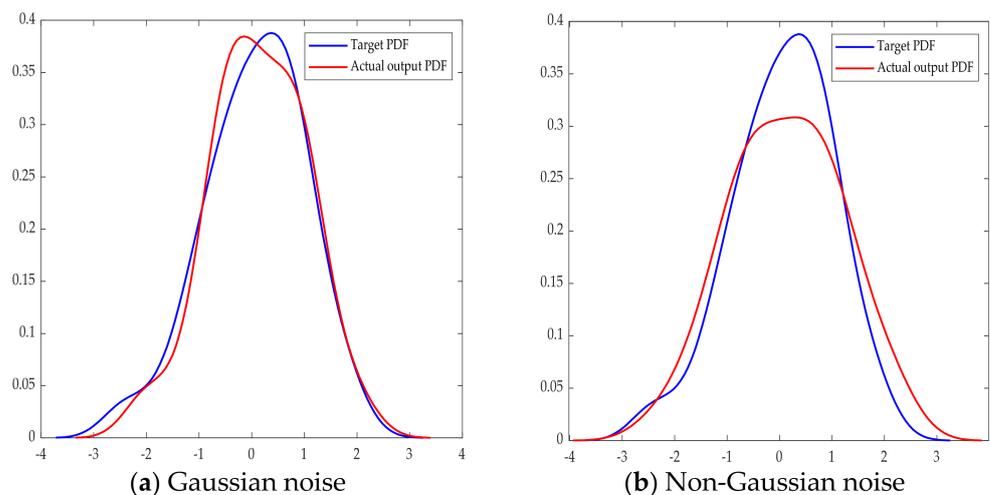
		N (0, 0.255)		B (2, 9)	
	Actual Value	Results of Traditional Correntropy	Results of Generalized Correntropy	Results of Traditional Correntropy	Results of Generalized Correntropy
$a_1$	-1	-0.9312	-0.9873	-1.0148	-1.0027
$b_0$	1	1.0061	0.9968	1.0254	1.0003
$b_1$	-1	-1.1073	-0.9816	-1.1995	-0.9992
$c_1$	-0.2	-0.1635	-0.1991	-0.1837	-0.2014
$F$	[1, 0.8]	[1, 0.7149]	[1, 0.7782]	[1, 0.8434]	[1, 0.8035]
$\hat{\eta}$		0.8566	0.8794	0.8927	0.9388
Time(s)		18.9443	15.2724	15.1667	13.6584
		E (0.5)		Bimodal Noise	
$a_1$	-1	-0.9838	-0.9884	-0.9945	-1.0097
$b_0$	1	0.9246	0.9961	1.0852	1.0643
$b_1$	-1	-0.9697	-0.9924	-1.0521	-1.0853
$c_1$	-0.2	-0.2194	-0.1935	-0.1798	-0.2037
$F$	[1, 0.8]	[1, 0.7520]	[1, 0.7857]	[1, 0.8102]	[1, 0.8138]
$\hat{\eta}$		0.9071	0.9399	0.9462	0.9336
Time(s)		18.9936	14.2616	18.5615	14.7962

The results show that when the disturbance obeys Gaussian, non-Gaussian and bimodal distribution, the identification results based on GCL are more accurate, and the time required to obtain the evaluation index is shorter.

4.2. Simulation When the Expected Output Distribution Is Known

Assumed that the known expected output distribution is Gaussian with mean value of 0 and variance of 1, as described in Section 2, the output PDF is expected to track the target PDF as much as possible.

The system under Gaussian and non-Gaussian noise disturbance are simulated respectively when the expected output distribution of the system is known. The output PDF tracks the input PDF as shown in Figure 7. The noise in (a) obeys normal distribution, while the noise in (b) obeys the non-Gaussian distribution. The generalized correntropy performance assessment index could be calculated by Formula (18) and shown in Table 5.



**Figure 7.** Expected and actual output PDF under different noises.

**Table 5.** Generalized correntropy performance index under expected output distribution is known.

	Index	Gaussian Noise	Non-Gaussian Noise
K = 0.8	$\eta_C$	0.8406	0.6554
	$\eta_{GC}$	0.8215	0.6021
K = 1.0	$\eta_C$	0.8698	0.7077
	$\eta_{GC}$	0.8520	0.6641
K = 1.2	$\eta_C$	0.8874	0.7447
	$\eta_{GC}$	0.8162	0.6429

In Table 5,  $\eta_{GC}$  is the generalized correntropy index and  $\eta_C$  is the traditional correntropy index. Assessment results of generalized correntropy are consistent with the traditional correntropy index when the controller gain  $K$  is different, which also proves the effectiveness of the new index.

## 5. Conclusions

A new CPA method for linear non-Gaussian systems based on generalized correntropy is proposed by analyzing the limitations of existing performance evaluation methods and indicators. The simple tracking index based on generalized correntropy is given directly when the expected output is known. When the expected output is unknown, the generalized correntropy loss function is used as the fitness value of the H-EDA to estimate the system parameters and noise distribution, so as to calculate the system performance benchmark more quickly and accurately, and then the performance assessment results can be obtained. The single variable method is used to simulate the SISO control system under different noise disturbances, and the effectiveness of the new index and H-EDA is verified respectively. The results show that generalized correntropy index is superior to the traditional correntropy in evaluating the control loop performance of linear systems with non-Gaussian stochastic disturbances, and the H-EDA could also identify the system parameters more quickly and accurately. Then, the follow-up research still focuses on the selection of the new performance assessment index and the improvement of identification algorithm.

**Author Contributions:** J.Z. and G.H. conceived and designed the experiments; G.H. performed the experiments; L.Z. and G.H. analyzed the data and wrote the paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the National Youth Science Foundation (Grant No. 61603136).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Jelali, M. *Control. Performance Management in Chemical Automation*; Springer: London, UK, 2013.
- Zhang, J.H.; Ren, M.F.; Tian, Y. Constrained stochastic distribution control for nonlinear stochastic systems with non-Gaussian noises. *Int. J. Innov. Comput. Inf. Control.* **2013**, *9*, 1759–1768.
- Harris, T.J. Assessment of Closed Loop Performance. *Can. J. Chem. Eng.* **1989**, *67*, 856–861. [[CrossRef](#)]
- Huang, B.; Shah, S.L. Performance Assessment of Nonminimum Phase Systems. In *Performance Assessment of Control Loops*; Springer: London, UK, 1999.
- Huang, B. Minimum variance control and performance assessment of time-variant processes. *J. Process. Control.* **2002**, *12*, 707–719. [[CrossRef](#)]
- Chen, J.; Wang, W.Y. Performance assessment of multivariable control systems using PCA control charts. In Proceedings of the IEEE Conference on Industrial Electronics and Applications (ICIEA 2009), Xi'an, China, 25–27 May 2009.
- Guo, L.; Wang, H. *Stochastic Distribution Control. System Design*; Springer: London, UK, 2010.
- Zhang, J.; Chu, C.C.; Munoz, J. Minimum entropy based on run-to-run control for semiconductor processes with uncertain metrology delay. *J. Process. Control.* **2009**, *19*, 1688–1697. [[CrossRef](#)]
- Zhang, J.; Jiang, M.; Chen, J. Minimum entropy-based performance assessment of feedback control loops subjected to non-Gaussian disturbances. *J. Process. Control.* **2014**, *24*, 1660–1670. [[CrossRef](#)]

10. Zhou, J.L.; Wang, X.; Zhang, J.F. A New Measure of Uncertainty and the Control Loop Performance Assessment for Output Stochastic Distribution Systems. *IEEE Trans. Autom. Control.* **2015**, *60*, 2524–2529. [[CrossRef](#)]
11. Zhou, J.; Jia, Y.; Jiang, H.; Fan, S. Non-Gaussian Systems Control Performance Assessment Based on Rational Entropy. *Entropy* **2018**, *20*, 331. [[CrossRef](#)]
12. Zhang, Q.; Wang, Y. Performance Assessment of Cascade Control System with Non-Gaussian Disturbance Based on Minimum Entropy. *Symmetry* **2019**, *11*, 379. [[CrossRef](#)]
13. Zhang, J.; Wu, D. Performance Assessment of Non-Gaussian Control Systems Based on Mixture Correntropy. *Entropy* **2019**, *21*, 1069. [[CrossRef](#)]
14. Zhang, Q.; Wang, Y. Improved Renyi Entropy Benchmark for Performance Assessment of Common Cascade Control System. *IEEE Access* **2019**, *7*, 6796–6803. [[CrossRef](#)]
15. Chen, B.; Wang, X.; Lu, N. Mixture Correntropy for Robust Learning. *Pattern Recognit.* **2018**, *79*, 318–327. [[CrossRef](#)]
16. Chen, B.; Xing, L.; Zhao, H. Generalized Correntropy for Robust Adaptive Filtering. *IEEE Trans. Signal. Process.* **2016**, *64*, 3376–3387. [[CrossRef](#)]
17. Sun, Q.; Zhang, T.; Huang, X. Robust Multi-kernel Generalized Maximum Correntropy Filters. In Proceedings of the 2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT), Shenzhen, China, 13–15 November 2020; pp. 99–103.
18. Wang, W.; Wang, S. Generalized Correntropy Sparse Gauss-Hermite Quadrature Filter for Epidemic Tracking on Complex Networks. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**. [[CrossRef](#)]
19. Ma, W.; Qiu, J.; Liu, X. Unscented Kalman Filter with Generalized Correntropy Loss for Robust Power System Forecasting-Aided State Estimation. *IEEE Trans. Ind. Inform.* **2019**, *15*, 6091–6100. [[CrossRef](#)]
20. Ren, M.; Gong, M.; Lin, M. Generalized Correntropy Predictive Control for Waste Heat Recovery Systems Based on Organic Rankine Cycle. *IEEE Access* **2019**, *7*, 151587–151594. [[CrossRef](#)]
21. Chen, B.; Zhu, Y.; Hu, J. *System Parameter Identification: Information Criteria and Algorithms*; Newnes: Boston, MA, USA, 2013.
22. Yu, J.; Zhou, J.; Li, D. Performance assessment of control loops with non-Gaussian disturbance Based on Generalized Minimum Entropy. In Proceedings of the International Conference on Industrial Artificial Intelligence (IAI), Shenyang, China, 23–27 July 2019.
23. Chen, B.; Principe, J.C. Maximum correntropy estimation is a smoothed MAP estimation. *IEEE Signal. Process. Lett.* **2012**, *19*, 491–494. [[CrossRef](#)]
24. Syed, M.N.; Pardalos, P.M.; Principe, J.C. On the optimization properties of the correntropic loss function in data analysis. *Optim. Lett.* **2014**, *8*, 823–839. [[CrossRef](#)]
25. Bjorklund, S. *A Survey and Comparison of Time-Delay Estimation Methods in Linear Systems*; Lund Institute of Technology: Lund, Sweden, 2003.
26. Chan, Y.T.; Wood, J. A new order determination technique for ARMA processes. *IEEE Trans. Acoust. Speech Signal. Process.* **1984**, *32*, 517–521. [[CrossRef](#)]
27. Doerr, B.; Krejca, M.S. Significance-Based Estimation-of-Distribution Algorithms. *IEEE Trans. Evol. Comput.* **2020**, *24*, 1025–1034. [[CrossRef](#)]
28. Wang, G.; Ma, L. The Estimation of Particle Swarm Distribution Algorithm with Sensitivity Analysis for Solving Nonlinear Bilevel Programming Problems. *IEEE Access* **2020**, *8*, 137133–137149. [[CrossRef](#)]
29. Liang, Y.; Ren, Z.; Yao, X. Enhancing Gaussian Estimation of Distribution Algorithm by Exploiting Evolution Direction with Archive. *IEEE Trans. Cybern.* **2020**, *50*, 140–152. [[CrossRef](#)] [[PubMed](#)]
30. Chi, Y.; Yao, Y.; Hu, L. Study of Via Optimization Problem Based on Estimation of Distribution Algorithm. In Proceedings of the International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 14–15 December 2019; pp. 84–88.
31. Gao, S.; Qiu, L.; Cao, C.G. Improved Estimation of Distribution Algorithm for Multi-objective Optimization Problems. *J. Nanjing Norm. Univ. (Nat. Sci. Ed.)* **2015**, *38*, 108–112.
32. Zhou, J.; Zhang, J.; Zhu, H. Non-Gaussian system identification based on improved estimation of distribution algorithm. In Proceedings of the 2017 36th Chinese Control Conference (CCC), Dalian, China, 26–28 July 2017; pp. 2052–2057.
33. Zhang, H.; Zhou, J.; Wang, J. Performance Assessment of Non-Gaussian Systems Based on Double Error Entropy Minimization. In Proceedings of the 2019 IEEE 8th Data Driven Control and Learning Systems Conference (DDCLS), Dali, China, 24–27 May 2019; pp. 1177–1182.