



## Article

# Diffusion Generalized MCC with a Variable Center Algorithm for Robust Distributed Estimation

Wentao Ma <sup>1,2,\*</sup> , Panfei Cai <sup>1</sup>, Fengyuan Sun <sup>2</sup>, Xiao Kou <sup>1</sup>, Xiaofei Wang <sup>3</sup> and Jianning Yin <sup>1</sup> 

- <sup>1</sup> School of Electrical and Engineering, Xi'an University of Technology, Xi'an 710048, China; 2201920038@stu.xaut.edu.cn (P.C.); 2190320002@stu.xaut.edu.cn (X.K.); yinjianning@xaut.edu.cn (J.Y.)
- <sup>2</sup> Guangxi Wireless Broadband Communication and Signal Processing Key Laboratory, Guilin University of Electronic Technology, Guilin 541004, China; fysun@guet.edu.cn
- <sup>3</sup> School of Microelectronics, Xi'an Jiaotong University, Xi'an 710049, China; mxfwang@xjtu.edu.cn
- \* Correspondence: mawt@xaut.edu.cn

**Abstract:** Classical adaptive filtering algorithms with a diffusion strategy under the mean square error (MSE) criterion can face difficulties in distributed estimation (DE) over networks in a complex noise environment, such as non-zero mean non-Gaussian noise, with the object of ensuring a robust performance. In order to overcome such limitations, this paper proposes a novel robust diffusion adaptive filtering algorithm, which is developed by using a variable center generalized maximum Correntropy criterion (GMCC-VC). Generalized Correntropy with a variable center is first defined by introducing a non-zero center to the original generalized Correntropy, which can be used as robust cost function, called GMCC-VC, for adaptive filtering algorithms. In order to improve the robustness of the traditional MSE-based DE algorithms, the GMCC-VC is used in a diffusion adaptive filter to design a novel robust DE method with the adapt-then-combine strategy. This can achieve outstanding steady-state performance under non-Gaussian noise environments because the GMCC-VC can match the distribution of the noise with that of non-zero mean non-Gaussian noise. The simulation results for distributed estimation under non-zero mean non-Gaussian noise cases demonstrate that the proposed diffusion GMCC-VC approach produces a more robustness and stable performance than some other comparable DE methods.

**Keywords:** diffusion adaptive filter; generalized Correntropy with variable center; non-zero mean non-Gaussian; distributed parameter estimation



check for updates

**Citation:** Ma, W.; Cai, P.; Sun, F.; Kou, X.; Wang, X.; Yin, J. Diffusion Generalized MCC with a Variable Center Algorithm for Robust Distributed Estimation. *Electronics* **2021**, *10*, 2807. <https://doi.org/10.3390/electronics10222807>

Academic Editor: Hirokazu Kobayashi

Received: 18 October 2021  
Accepted: 15 November 2021  
Published: 16 November 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/>).

## 1. Introduction

Distributed estimation has become an important technology. Its object is to estimate interesting and available parameters from noisy measurements using a cooperation strategy between nodes over networks for distributed network applications, such as environment monitoring, spectrum sensing, and source localization [1–3]. In recent years, diffusion adaptive filtering (DAF) algorithms with the mean square error (MSE) criterion have been proven to be an optimal and effective method for distributed estimation (DE) in additive white Gaussian noise environments. Among all DAFs, the diffusion least mean square (DLMS) [4–6] and diffusion recursive least square [7] are outstanding representatives that have received significant attention in DE applications. However, the performance of these methods with the MSE criterion may degrade under non-Gaussian environments, such as lightning noise, sea cluttering noise, and co-channel interference in distributed spectrum sensing for Cognitive radio applications [8].

Recently, an increasing number of researchers have focused on the development of robust DAF algorithms via the non-second order statistical method. For this purpose, a robust DE method based on the mean p-Power error criterion, called diffusion least mean p-power (DLMP), was developed to estimate the parameters of wireless sensor networks [9]. Specially, the diffusion least mean fourth [10,11] and the diffusion sign

error-LMS(DSE-LMS) [12] algorithms, as special cases of the DLMP, were proposed for DE over networks in the cases of non-Gaussian noise interference. In addition, the maximum Correntropy criterion (MCC) was defined to extend second-order statistics into higher-order statistics via the exploitation of the Gaussian kernel function, which can be used as a cost function to design robust adaptive filters due to its smoothness and strict positive-definiteness [13,14]. A robust DE method based on MCC, called diffusion MCC (DMCC), was developed in [15] to mitigate the robustness of the traditional DAF algorithms. In addition, a proportional DMCC algorithm with adaptable kernel width was proposed for sparse distributed system identification in the cases of impulse noise [16]. The diffusion subband adaptive filtering (DSAF) algorithm, based on symmetrical MCC with individual weighting factors, was developed for colored input signals [17]. To improve the convergence performance of the conventional diffusion affine projection (AP) algorithm, an MCC-based diffusion AP algorithm was further derived using the MCC as the cost function for DE over networks [18]. However, the Gaussian kernel in Correntropy may not always be an ideal option under some specific conditions [19]. Consequently, a novel generalized form of Correntropy was further defined by Chen [19], in which the generalized Gaussian density (GGD) function is utilized as the kernel function in Correntropy [19]. Similarly, generalized Correntropy can also be used as a cost function in adaptive signal processing and machine learning fields, and is called generalized MCC (GMCC). The GMCC may achieve better performance than the MCC-based methods for measurements in non-Gaussian noise environments [20–22]. This is because a greater number of higher-order moments of data containing errors are contained in GMCC and the additional shape parameters introduced by GGD can further expand the range of possible induced metrics. Therefore, GMCC has been widely utilized to design various robust adaptive filters, such as kernel adaptive filtering under GMCC [22], kernel recursive GMCC [20], Stacked Extreme Learning Machine (ELM) with GMCC [23], and the unscented Kalman filter with GMCC [24]. Chen et al. proposed the diffusion GMCC method for distributed estimation [25], and a novel robust diffusion affine projection GMCC algorithm was further developed over networks [26]. Although the GMCC-based methods can achieve good performance in non-Gaussian noise cases, the steady-state performance may degrade in some practical situations because the center of the generalized Gaussian kernel is located at zero. For example, the error criteria located at zero cannot obtain outstanding results when the error distribution of the signal has a non-zero mean. The main reason for this is that the zero-mean Gaussian function usually cannot match the error distribution well in this case. To overcome this problem, a variable center was introduced into the MCC to define a novel MCC-based criterion [27], called MCC-VC. The MCC-VC-based adaptive filtering methods can achieve better performance than the original MCC-based methods under cases of non-zero mean non-Gaussian noise because of the non-zero center. Taking advantage of the MCC-VC, several MCC-VC-based adaptive filtering algorithms [28,29] and ELM with MCC-VC [30] have been proposed for signal processing and machine learning applications.

Inspired by the MCC-VC and considering the property of the GMCC, a GMCC with a variable center (GMCC-VC) was defined by the author [30], and a recursive adaptive filtering algorithm with a sparse penalty term based on GMCC-VC was developed for sparse system estimation under non-zero mean non-Gaussian environments. In this paper, we focus on the development of a novel robust diffusion adaptive filtering algorithm based on the GMCC-VC, because the center can be located anywhere to obtain good performance for DE over network in more common situations. Due to the attribute of the insensitivity to the outliers, especially with a small kernel bandwidth, the use of GMCC is able to further mitigate the negative impact of non-Gaussian (impulsive) noise on the estimation performance. Moreover, the variable center strategy is used to locate the center anywhere in order to improve the robustness of the proposed method in non-zero mean non-Gaussian noise environments. For feasibility, the online parameter optimization method was designed using the gradient approach to improve the performance of the proposed algorithm. Simulation results demonstrate that the proposed method can effectively

improve the distributed parameter estimation over networks in the presence of non-zero mean non-Gaussian noise.

The remainder of this paper is organized as follows. In Section 2, we briefly review generalized Correntropy and define the GMCC with a variable center. In Section 3, the diffusion GMCC with the variable center algorithm is developed and the parameters' optimization methods are presented. In Section 4, numerical simulations are performed to test the performance of the proposed algorithm results. Finally, we conclude this work in Section 5.

## 2. Generalized Maximum Correntropy Criterion with Variable Center

In order to improve the performance of the original generalized Correntropy in the cases of non-zero mean non-Gaussian noise, we introduce the variable center idea to generalized Correntropy to extend its applications. First, this section briefly reviews generalized Correntropy, and then defines generalized Correntropy with a variable center.

### 2.1. Briefly Review of the Generalized Correntropy

Generalized Correntropy with a GGD function between arbitrarily given random variables  $X$  and  $Y$  can be defined as [19]:

$$\mathbf{V}_{\alpha,\tau}(X, Y) = \mathbf{E}(\mathbf{G}_{\alpha,\tau}(X - Y)) \quad (1)$$

where  $\mathbf{E}[\cdot]$  denotes the expectation operator. The GGD function with a zero-mean is usually used as the kernel function in Equation (1), which is expressed in the following form:

$$\begin{aligned} \mathbf{G}_{\alpha,\tau}(e) &= \frac{\alpha}{2\beta\Gamma(1/\alpha)} \exp\left(-\left|\frac{e}{\beta}\right|^\alpha\right) \\ &= \gamma_{\alpha,\tau} \exp(-\tau|e|^\alpha) \end{aligned} \quad (2)$$

where  $e = x - y$ ,  $\alpha > 0$  denotes the shape parameter,  $\beta > 0$  represents the bandwidth parameter,  $\tau = \frac{1}{\beta^\alpha}$  is the kernel parameter, and  $\gamma_{\alpha,\tau} = \frac{\alpha}{2\beta\Gamma(1/\alpha)}$  stands for a normalized constant. The GGD in Equation (2) is more general and flexible, and general Correntropy may achieve good capability for complex noise cases. In addition, Correntropy with the Gaussian kernel is a special case when generalized Correntropy is set using suitable parameters.

In general, it is difficult to know the joint distribution of  $X$  and  $Y$ , and only finite number of samples  $\{(x_i, y_i)\}_{i=1}^N$  are available. Therefore, the sample mean estimator of generalized Correntropy is usually defined and used in practice, and is expressed as:

$$\hat{\mathbf{V}}_{\alpha,\tau}(X, Y) = \frac{1}{N} \sum_{i=1}^N \mathbf{G}_{\alpha,\tau}(x_i - y_i) \quad (3)$$

where  $N$  is the number of sample points. The generalized Correntropy of error can be utilized as a cost function to design robust adaptive filtering algorithms. This is called the generalized maximum Correntropy criterion (denoted as GMCC).

Furthermore, a generalized C-loss function can be defined as:

$$\begin{aligned} J_{GC-loss}(X, Y) &= \gamma_{\alpha,\tau} - \mathbf{V}_{\alpha,\tau}(X, Y) \\ &= \gamma_{\alpha,\tau} - \mathbf{E}(\mathbf{G}_{\alpha,\tau}(X - Y)) \\ &= \gamma_{\alpha,\tau} \{1 - \mathbf{E}(\exp(-\tau|e|^\alpha))\} \end{aligned} \quad (4)$$

From Equation (3), one can easily obtain the sample mean estimator of the generalized C-loss in Equation (4) as:

$$\begin{aligned} \widehat{J}_{GC-loss}(X, Y) &= \gamma_{\alpha, \tau} - \frac{1}{N} \sum_{i=1}^N \mathbf{G}_{\alpha, \tau}(x_i - y_i) \\ &= \gamma_{\alpha, \tau} \left( 1 - \frac{1}{N} \sum_{i=1}^N \exp(-\tau |e_i|^\alpha) \right) \end{aligned} \tag{5}$$

One can easily find that minimizing  $J_{GC-loss}(X, Y)$  is equivalent to maximize  $\mathbf{V}_{\alpha, \tau}(X, Y)$ .

### 2.2. Generalized Maximum Correntropy Criterion with Variable Center

As mentioned in [19], generalized Correntropy with the GGD kernel can achieve good performance, and now many generalized C-loss based adaptive filtering and machine learning methods have been developed in different applications. However, the performance of GMCC with zero mean GGD may degrade due to interference with a non-zero mean noise distribution. Therefore, it is important and of interest to expand the flexibility of generalized Correntropy so that it can be adapted to the aforementioned situations.

Inspired by Correntropy with a variable center, we define generalized Correntropy with a variable center (GC-VC) between  $X$  and  $Y$  as [31]:

$$\begin{aligned} \mathbf{V}_{\tau, c}(X, Y) &= \mathbf{E}[\mathbf{G}_{\tau, c}(x - y - c)] \\ &= \mathbf{E}[\gamma_{\alpha, \tau} \exp(-\tau |x - y - c|^\alpha)] \\ &= \mathbf{E}[\gamma_{\alpha, \tau} \exp(-\tau |e - c|^\alpha)] \end{aligned} \tag{6}$$

where the center location  $c \in \mathfrak{R}$  should be optimized, and mainly controls the performance of the GC-VC. By comparing Equations (1) and (6), it can be seen that the GC-VC will reduce to generalized Correntropy when the center is set at zero, and Correntropy with a variable center when  $\alpha = 2$ . The GC-VC also involves higher order moments of the error about the center  $c$  as:

$$\mathbf{V}_{\tau, c}(X, Y) = \gamma_{\alpha, \tau} \sum_{n=0}^{\infty} \frac{(-\tau)^n}{n!} \mathbf{E}[|e - c|^{\alpha n}] \tag{7}$$

Similar to generalized Correntropy, the sample estimator of the GC-VC can be given as:

$$\widehat{\mathbf{V}}_{\alpha, \tau}(X, Y) = \frac{1}{N} \sum_{i=1}^N \mathbf{G}_{\tau, c}(x_i - y_i - c) \tag{8}$$

Furthermore, one can easily obtain the GC-VC loss as:

$$\begin{aligned} J_{GC-VC-loss}(X, Y) &= \gamma_{\alpha, \tau} - \mathbf{V}_{\tau, c}(X, Y) \\ &= \gamma_{\alpha, \tau} - \mathbf{E}[\mathbf{G}_{\tau, c}(X - Y - c)] \\ &= \gamma_{\alpha, \tau} \{ 1 - \mathbf{E}(\exp(-\tau |e - c|^\alpha)) \} \end{aligned} \tag{9}$$

Now, an optimal model under the GC-VC (or GC-VC loss) can be obtained as:

$$\begin{aligned} \mathbf{M}^* &= \underset{M \in \wp}{\operatorname{argmax}} \mathbf{V}_{\tau, c}(x, y) = \mathbf{E}[\mathbf{G}_{\tau, c}(e - c)] \\ &\Leftrightarrow \underset{M \in \wp}{\operatorname{argmin}} J_{GC-VC-loss}(x, y) = \gamma_{\alpha, \tau} \{ 1 - \mathbf{E}(\exp(-\tau |e - c|^\alpha)) \} \end{aligned} \tag{10}$$

where  $\mathbf{M}^*$  denotes the optimal model, and  $\wp$  is the models' hypothesis space. We then call the optimal model GMCC with a variable center (GMCC-VC).

### 3. Diffusion Adaptive Filtering Algorithm under GMCC-VC

Diffusion adaptive filtering (DAF) algorithms have been widely used for distributed estimation over networks due to their outstanding performance. However, the tradi-

tional DAF algorithms under MSE cannot achieve a good performance in cases of complex non-Gaussian noise. In this section, we develop a novel robust DAF under the GMCC-VC criterion.

### 3.1. Signal Model and Diffusion GMCC-VC

Here, we consider a network model that is composed of  $N$  nodes distributed over a geographic area to estimate an unknown vector  $\mathbf{w}_o$  of size  $(M \times 1)$  from measurements collected at  $N$  nodes. Here, each node  $k$  can access the realization of a scalar measurement  $d_k(i)$  and a regression vector  $\mathbf{u}_k(i)$  of size  $(M \times 1)$  at each time index  $i$  ( $i = 1, 2, \dots, I$ ), related as:

$$d_k(i) = \mathbf{w}_o^T \mathbf{u}_k(i) + v_k(i) \tag{11}$$

where  $v_k(i)$  represents the measurement noise and  $T$  stands for transposition.

Based on the model mentioned above, we develop the diffusion GMCC-VC (DGMCC-VC) algorithm for each node  $k$  to estimate  $\mathbf{w}_o$  by maximizing a linear combination of local generalized Correntropy with a variable center within the node  $k$ 's neighbor  $N_k$ . Then, we define the cost function of the DGMCC-VC for each node  $k$  as:

$$\begin{aligned} J_k^{local}(w_k) &= \sum_{l \in N_k} \delta_{l,k} \mathbf{G}_{\tau,c}(e_{l,k}(i)) \\ &= \sum_{l \in N_k} \delta_{l,k} \mathbf{G}_{\tau,c}(d_l(i) - \mathbf{w}_k^T(i) \mathbf{u}_l(i) - c_l) \end{aligned} \tag{12}$$

where  $\mathbf{w}_k(i)$  denotes the estimate of node  $k$  for  $\mathbf{w}_o$  at time instant  $i$ ,  $e_{l,k}(i) = d_l(i) - \mathbf{w}_k^T(i) \mathbf{u}_l(i) - c_l$  is estimate error,  $\{\delta_{l,k}\}$  are some non-negative combination coefficients, and

$$\begin{aligned} \mathbf{G}_{\tau,c}(e_{l,k}(i)) &= \gamma_{\alpha,\tau} \exp\left(-\tau |e_{l,k}(i) - c_l|^\alpha\right) \\ &= \gamma_{\alpha,\tau} \exp\left(-\tau |d_l(i) - \mathbf{w}_k^T \mathbf{u}_l(i) - c_l|^\alpha\right) \end{aligned} \tag{13}$$

In general, the adapt-then-combine (ATC) strategy is usually used to design the diffusion adaptive filtering algorithm because it can achieve lower steady-state misalignment compared with the combine-then-adapt (CTA) diffusion strategy in some situations [5]. As a result, we mainly focus on the development of the ATC diffusion GMCC-VC (briefly denoted as DGMCC-VC) algorithm in this work, which can be given by the following adaptation and combination steps as:

$$\begin{aligned} \psi_k(i) &= \mathbf{w}_k(i-1) + \mu_k \nabla J_k^{local}(\mathbf{w}_k) \\ &= \mathbf{w}_k(i-1) + \mu_k \sum_{l \in N_k} \delta_{l,k} \mathbf{G}_{\tau,c}(e_{l,k}(i)) |e_{l,k}(i) - c_l|^{\alpha-1} \text{sign}(e_{l,k}(i) - c_l) \mathbf{u}_l(i) \end{aligned} \tag{14}$$

$$\mathbf{w}_k(i) = \sum_{l \in N_k} \beta_{l,k} \psi_l(i) \tag{15}$$

where  $\psi_k(i)$  represents an intermediate estimate for  $\mathbf{w}_o$  supplied by node  $k$  at time instant  $i$ .  $\beta_{l,k}$  denotes the combination weight of agent  $l$  on agent  $k$ . Generally,  $\delta_{l,k}$  and  $\beta_{l,k}$  should be set to satisfy the following conditions:

$$\begin{aligned} \delta_{l,k} \geq 0, \sum_{k=1}^N \delta_{l,k} = 1, \text{ and } \delta_{l,k} = 0 \quad \text{if } l \notin N_k \\ \beta_{l,k} \geq 0, \sum_{k=1}^N \beta_{l,k} = 1, \text{ and } \beta_{l,k} = 0 \quad \text{if } l \notin N_k \end{aligned} \tag{16}$$

Recently, several rules have been instituted for selecting these weights, such as the uniform, Metropolis, maximum degree, relative degree, and relative degree-variance rules; their detailed methods can be viewed in [4].

### 3.2. Free Parameter Optimization

Equation (13) shows that two free parameters (the center and the kernel width) are contained in the DGMCC-VC algorithm, which have significant influence on its performance. Optimize these parameters is the crucial problem for this algorithm. In this subsection, we use an online parameter adaptation approach to optimize these parameters; the optimal model is as follows:

$$(\tau_l(i), c_l(i)) = \underset{\tau \in T, c \in \mathbb{C}}{\operatorname{argmin}} \{ \gamma_{\tau, c} - \mathbf{E}[\mathbf{G}_{\tau, c}(e_l(i) - c_l)] \} \tag{17}$$

where the admissible sets of parameters  $\tau_l$  and  $c_l$  are represented as  $T$  and  $\mathbb{C}$ ,  $\tau_l(i)$  and  $c_l(i)$  denote the adapted parameters at iteration time  $i$ . In general, using the Parzen window theory, we have the following results as  $\mathbf{E}[\mathbf{G}_{\tau_l, c_l}(e_l(i) - c_l)] \triangleq \sum_{i=n-L}^n \mathbf{G}_{\tau_l, c_l}(e_l(i) - c_l)$ , where the window length of the error samples is  $L$ . From [17], it can be seen that the gradient-based approach can be used to solve the optimization problem in Equation (11) in a given finite set. In this work, the following simple methods are used iteratively to optimize the free parameters  $c_l$  and  $\tau_l$ :

Center  $c$ : The mean or median value of the error samples are used to obtain the estimate of the center parameter  $c$ , and the method can be given as:

$$c_l(i) = \operatorname{median}\{\operatorname{sort}\{|e_l(i)|, |e_l(i-1)|, \dots, |e_l(i-L+1)|\}\} \tag{18}$$

where the window length  $L$  is usually selected to ensure the fit of the error curve to estimate the parameters [31],  $\operatorname{sort}\{X\}$  denotes a sort function which sorts the elements of the vector  $X$  in ascending order, and  $\operatorname{median}\{X\}$  represents the median value of the elements in  $X$ .

Kernel width  $\tau$ : To select an optimal kernel width, we use the gradient based method to adaptively optimize this free parameter at each iteration. Taking the derivative of Equation (3) with respect to  $\tau$ , we can formulate a simple gradient descent-based search rule to update the kernel width as:

$$\begin{aligned} \tau_l(i+1) &= \tau_l(i) - \mu_\tau \frac{\partial}{\partial \tau_l(i)} \{ \gamma_{\alpha, \tau} \exp(-\tau_l(i)|e_l(i) - c_l(i)|^\alpha) \} \\ &= \tau_l(i) + \mu_\tau \gamma_{\alpha, \tau} |e_l(i) - c_l(i)|^\alpha \exp(-\tau_l(i)|e_l(i) - c_l(i)|^\alpha) \\ &= \tau_l(i) + \eta_\tau \exp(-\tau_l(i)|e_l(i) - c_l(i)|^\alpha) |e_l(i) - c_l(i)|^\alpha \end{aligned} \tag{19}$$

where  $\eta_\tau = \mu_\tau \gamma_{\alpha, \tau}$  denotes the step-size parameter for update of  $\tau_l$ .

### 3.3. DGMCC-VC Algorithm with No Measurement Exchange

Equation (14) shows that the exchange of data is required during the adaptation stage, which will make the communication cost relatively high. To address this problem, the uncomplicated strategy with no measurement exchange is used in adaptation stage, and then, using the parameters optimization method, we can obtain the updated equations of the adaptation and combination stage of the novel DGMCC-VC algorithm as:

$$\begin{cases} \psi_k(i) = \mathbf{w}_k(i-1) + \eta_k [G_{\tau, c}(e_k^c(i)) |e_k^c(i)|_{\alpha-1} \operatorname{sign}(e_k^c(i)) \mathbf{u}_k(i)] \\ \mathbf{w}_k(i) = \sum_{l \in N_k} \delta_{l, k} \psi_l(i) \end{cases} \tag{20}$$

where  $e_k^c(i) = d_k(i) - \mathbf{u}_k^T(i) \mathbf{w}_k(i-1) - c_k(i)$  is the extended error with a variable center for node  $k$ . The DGMCC-VC algorithm is summarized schematically in Algorithm 1.



**Algorithm 1** Diffusion GMCC-VC

---

Initialized coefficients  $\{\alpha, \eta_k\}$   
 Start with  $\mathbf{w}_k(-1) = 0$ .  
**for**  $i = 0$  to  $I$  **do**  
     **for**  $k = 1$  to  $N$  **do**  
          $e_k^c(i) = d_k(i) - \mathbf{u}_k(i)\mathbf{w}_k(i-1) - c_k(i)$   
         The free parameters  $\tau$  and  $c$  optimization according to Equations (18) and (19)  
          $G_{\tau,c}(e_k^c(i)) = \exp(-\tau_k(i)|e_k^c(i)|^\alpha)$   
          $\psi_k(i) = \mathbf{w}_k(i-1) + \eta_k[G_{\tau,c}(e_k^c(i))|e_k^c(i)|_{\alpha-1}\text{sign}(e_k^c(i))\mathbf{u}_k(i)]$   
          $\mathbf{w}_k(i) = \sum_{l \in N_k} \delta_{l,k}\psi_l(i)$   
     **end for**  
**end for**

---

**Remark:** An extra exponential function of the error  $G_{\tau,c}(e_k^c(i))$  introduced by the GMCC-VC is contained in Equation (20), and this scaling factor will approach zero when a large error occurs (possibly caused by an outlier), which endows the DGMCC-VC algorithm with the property of resisting the influence of outliers. The DGMCC-VC algorithm can be viewed as the DGMCC algorithm when we set the center at zero. In addition, the DGMCC algorithm will reduce to the DLMP algorithm [9] when  $G_{\tau,c}(e_k^c(i))$  is one and the center is located at zero. In addition, it can easily be found that the DLMS, DLMF, and DSE-LMS are special cases of the DGMCC-VC algorithm.

#### 4. Simulation Results

In this section, we perform Monte Carlo (MC) simulations to verify the performance of the proposed DGMCC-VC algorithm for distributed parameter estimation over networks in non-Gaussian and non-zero mean noise environments. Here, we consider a network topology with 20 nodes, which is generated as a realization of the random geometric graph model, and the unknown parameter vector is set to  $\frac{\text{randn}(M,1)}{\sqrt{M}}$  ( $M = 10$ ), where  $\text{randn}(\cdot)$  represents a function that generates random values with a Gaussian distribution. The input signals are assumed to be zero-mean Gaussian with size  $M = 10$ . All results are calculated by taking the ensemble average of the network MSD over 200 independent MC runs. Furthermore, the linear combination coefficients are obtained using the Metropolis rule [4]. In particular, the performance of the proposed DGMCC-VC is compared with some existing algorithms, including the DLMS, DMCC [15], DLMP [9], DSE-LMS [12], and DGMCC algorithms. In order to test the convergence and steady-state performance, we define the mean square deviation (MSD) given in Equation (21) as the evaluation criterion:

$$MSD = 10 \log \left[ \frac{1}{I} \sum_{i=1}^I \|\mathbf{w}_o - \mathbf{w}(i)\|^2 \right] \quad (21)$$

The measurement noise  $v(i)$  is composed of two independent noises, which can be expressed as  $v(i) = (1 - p(i))A(i) + p(i)B(i)$ , where  $A(i)$  with a non-zero mean is inner noise,  $B(i)$  with much larger variance is used to model outliers, and  $p(i)$  denotes an independent and identically distributed binary process with an occurrence probability of 0.05. In the following simulations, we assume that the noises  $A(i)$  and  $B(i)$  are independent of  $a(i)$ , and  $B(i)$  is a white Gaussian process with a mean of zero and variance of 100. The following three non-zero mean non-Gaussian distributions are considered as inner noise  $A(i)$ , which follows [31]:

- (1) Uniform distribution over [1,2];
- (2) Laplace distribution with mean of one and unit variance;
- (3) Binary distribution over {0,1} with probability mass  $p[x = -1] = p[x = 1] = 0.5$ .

#### 4.1. Performance Comparison among the Proposed Algorithm and Other Algorithms

We investigated the steady-state performance of the proposed algorithm in different non-mean non-Gaussian noise environments. For each simulation, the number of repetitions was set at 1000. The step size values of all algorithms were set to ensure almost the same initial convergence rate, which can be seen in the legend. The  $p$  was set at 1.1 for the DLMP algorithm, the kernel size was selected as 2.0 for the DMCC algorithm, and the kernel width for DGMCC was 0.1. The exponent parameters were set to 1.8, 1.8, and 2.5 for GMCC and GMCC-VC algorithms for the different inner noises mentioned above, respectively. All parameters were set by scanning for the best results. The convergence curves in terms of MSD are shown in Figure 1a, Figure 2a, Figure 3a under the inner noises (1)–(3), respectively. We can clearly see that the proposed DGMCC-VC algorithm outperforms other methods in terms of steady-state accuracy. The results confirm that the proposed algorithm exhibits a significant improvement in steady-state performance in non-zero mean non-Gaussian noise environments due to the variable center strategy. Furthermore, the steady-state MSDs at each node  $k$  are given in Figure 1b, Figure 2b, Figure 3b, respectively. These figures demonstrate the above conclusion that the DGMCC-VC algorithm still shows good performance compared with all other algorithms.

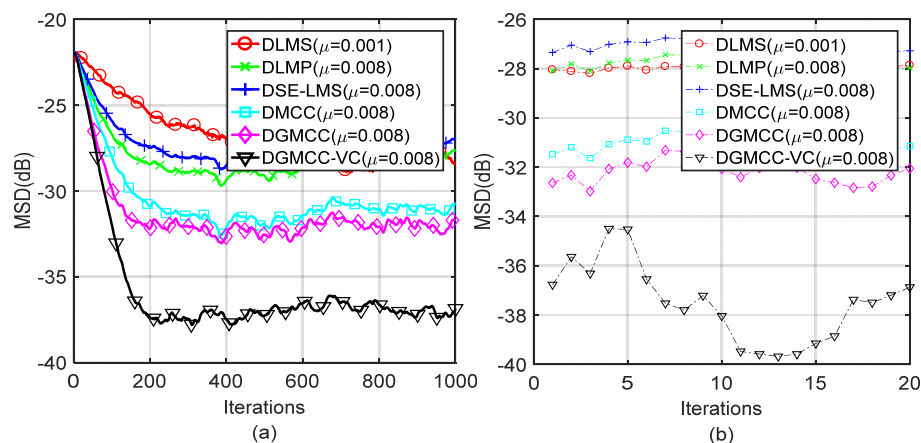


Figure 1. Results under the inner noise with uniform distribution: (a) convergence curves in terms of MSD; (b) MSD at steady state for 20 nodes.

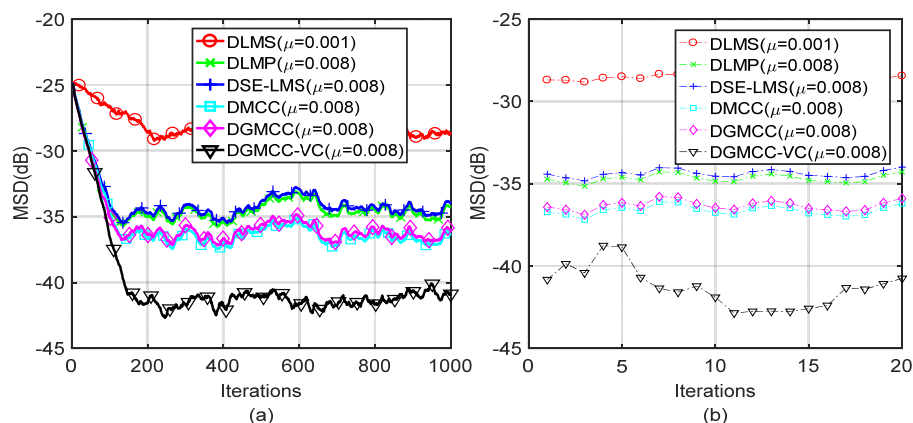
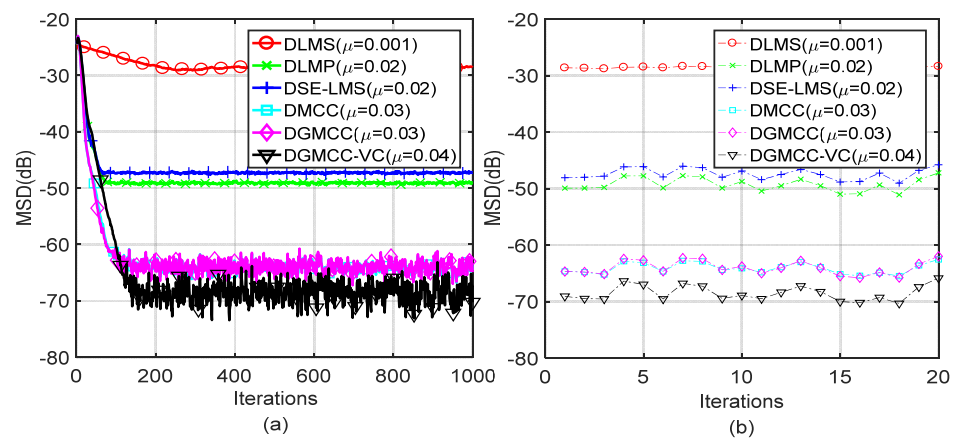


Figure 2. Results under the inner noise with Laplace distribution: (a) convergence curves in terms of MSD; (b) MSD at steady state for 20 nodes.

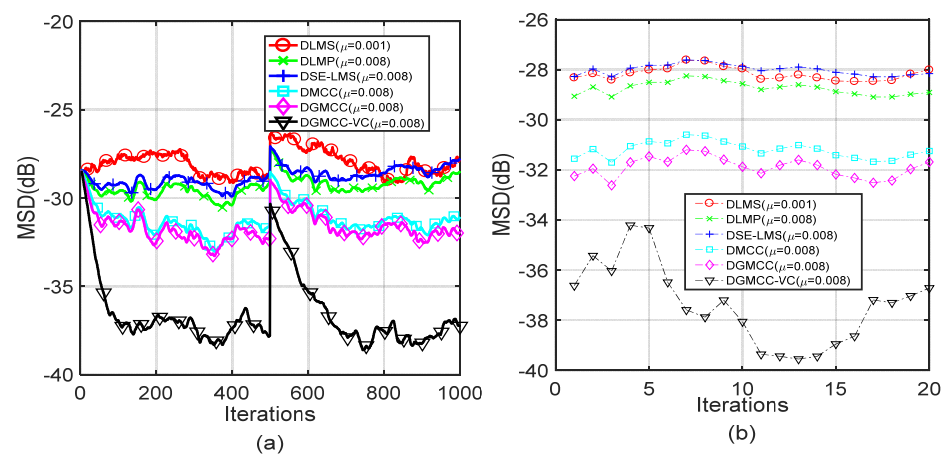




**Figure 3.** Results under the inner noise with binary distribution: (a) convergence curves in terms of MSD; (b) MSD at steady state for 20 nodes.

#### 4.2. Performance Comparison under Time-Varying Parameter Estimation

In order to test the tracking capabilities of the proposed method, we considered a time-varying parameter estimation case in which the unknown system is changed at the middle of the iterations. Here the number of the iterations is 1000. The parameters of the unknown system are set to different values before and after 500 iterations. Here the inner noise follows a uniform distribution. The convergence curves and MSD at steady state are shown in Figure 4. The results in Figure 4 show that: (1) the proposed algorithm achieves better steady-state accuracy at different stages compared with other methods; (2) the algorithm can converge quickly when the unknown parameters are changed, which means that the proposed algorithm has good tracking ability.



**Figure 4.** Results for the estimation of time-varying parameters under the condition of inner noise with uniform distribution: (a) convergence curves in terms of MSD; (b) MSD at steady state for 20 nodes.

#### 4.3. Performance of the Proposed Algorithm with Different Free Parameters

From Equations (18) and (19), we know that the window length  $L$  and step size are used to adaptively optimize the center and kernel width. In this subsection, we further investigate the effect of these two free parameters on the performance of the proposed DGMCC-VC algorithm under the condition of uniform inner noise, as represented in Equation (1). First, we set the  $L$  at 5, 15, 20, 25, and 30. The other simulation settings were consistent with those of the previous simulations. The convergence curves of the proposed algorithm with different  $L$  values are plotted in Figure 5. We can observe that the proposed DGMCC-VC can converge under all selected values of  $L$ , but a good performance

is obtained when  $L = 5$  in this case. Second, we performed simulations to investigate the performance of the DGMCC-VC algorithm with different step size values of 0.05, 0.08, 0.1, 0.12, 0.15, and 0.20. The results in Figure 6 show that the proposed method can converge consistently for different step size values, and the performance steadily increases when the step value gradually increases from 0.05 to 0.2. According to the results above, we conclude that the free parameters in the optimal method for the center and kernel width are still important for the proposed method.

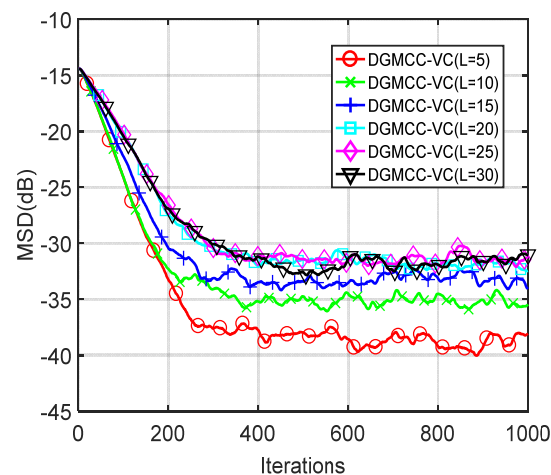


Figure 5. Convergence curves of the DGMCC-VC algorithm with different  $L$ .

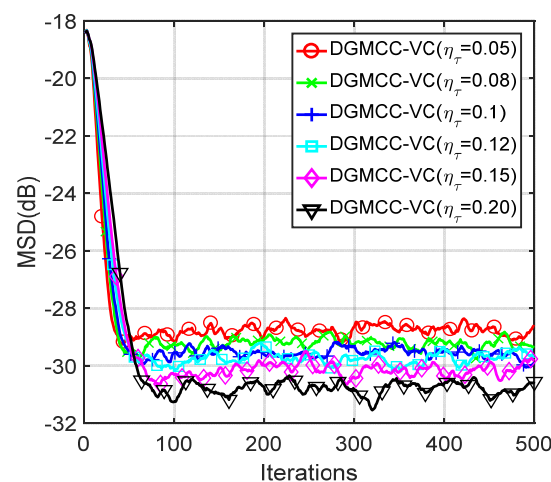


Figure 6. Convergence curves of DGMCC-VC algorithm with different  $\eta_\tau$ .

## 5. Conclusions

This paper proposed a novel diffusion adaptive filter (DAF) based on generalized maximum Correntropy with a variable center (GMCC-VC) to improve the performance of classical DAFs for distributed estimation over a network in a non-zero mean non-Gaussian noise environment. Generalized Correntropy with a variable center via the generalized Gaussian kernel function was defined to match the non-zero mean distribution of the non-Gaussian noise. Then, a novel robust diffusion adaptive filtering algorithm based on the GMCC-VC was designed using the adapt-then-combine strategy for distributed estimation over networks. The free parameter optimization techniques based on the gradient method were employed to improve the performance of the proposed algorithm. Simulation results demonstrate that the proposed method outperforms the existing comparable methods for distributed estimation in the case of non-zero mean non-Gaussian noise.

Although the proposed method shows outstanding performance for distributed estimation over networks in special cases, some limitations remain, namely: (1) how to adaptively select the optimal shape parameter  $\alpha$  under different conditions; and (2) how to reduce the time complexity of the parameters' optimal process. Those two limitations may be challenges and directions for our future research. Furthermore, sparse distributed estimation in the case of non-zero mean non-Gaussian noise will also be a meaningful area of study in the future.

**Author Contributions:** Conceptualization, W.M.; methodology, W.M. and F.S.; software, P.C. and X.K.; validation, W.M. and J.Y.; formal analysis, W.M. and J.Y.; investigation, W.M. and X.K.; writing—original draft preparation, W.M. and P.C.; writing—review and editing, W.M. and X.W.; project administration, W.M. and X.W.; funding acquisition, W.M. and F.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China under Grant 6197617 and 52107167, the Key Project of Natural Science Basic Research Plan in Shaanxi Province of China under Grant 2019JZ-05, the Key Laboratory Project of Shaanxi Provincial Education Department Scientific Research Projects under Grant 20JS109, the Dean Project of Guangxi Wireless Broadband Communication and Signal Processing Key Laboratory, the Basic research program of Natural Science in Shaanxi Province under Grant 2021JQ-473, the Basic research program of Natural Science in Shaanxi Province under Grant 2021JQ-473, and the China Postdoctoral Science Foundation under Grant 2021M692877.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

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

## References

1. Huang, W.; Yang, X.; Liu, D.; Chen, S. Diffusion LMS with component-wise variable step-size over sensor networks. *IET Signal Process.* **2016**, *10*, 37–45. [[CrossRef](#)]
2. Meng, W.; Xiao, W.; Xie, L. A Projection Based Fully Distributed Approach for Source Localization in Wireless Sensor Networks. *Ad. Hoc. Sens. Wirel. Netw.* **2013**, *18*, 131–158.
3. Ainomäe, A.; Bengtsson, M.; Trump, T. Distributed largest eigenvalue-based spectrum sensing using diffusion LMS. *IEEE Trans. Signal Inf. Process. Over Netw.* **2018**, *4*, 362–377. [[CrossRef](#)]
4. Cattivelli, F.; Sayed, A.H. Diffusion LMS strategies for distributed estimation. *IEEE Trans. Signal Processing* **2010**, *58*, 1035–1048. [[CrossRef](#)]
5. Lopes, C.G.; Sayed, A.H. Diffusion least-mean squares over adaptive networks: Formulation and performance analysis. *IEEE Trans. Signal Process.* **2008**, *56*, 3122–3136. [[CrossRef](#)]
6. Azam, K.; Mohammad, A.T.; Amir, R.; Jonathon, A.C. Transient analysis of diffusion least-mean squares adaptive networks with noisy channels. *Int. J. Adapt. Control Signal Process.* **2012**, *26*, 171–180.
7. Cattivelli, F.S.; Lopes, C.G.; Sayed, A.H. Diffusion recursive least-squares for distributed estimation over adaptive networks. *IEEE Trans. Signal Process.* **2008**, *56*, 1865–1877. [[CrossRef](#)]
8. Liu, M.; Zhao, N.; Li, J.; Leung, V.C.M. Spectrum Sensing Based on Maximum Generalized Correntropy Under Symmetric Alpha Stable Noise. *IEEE Trans. Veh. Technol.* **2019**, *68*, 10262–10266. [[CrossRef](#)]
9. Wen, F. Diffusion least-mean P-power algorithms for distributed estimation in alpha-stable noise environments. *Electron Lett.* **2013**, *49*, 1355–1356. [[CrossRef](#)]
10. Ni, J.; Yang, J. Variable step-size diffusion least mean fourth algorithm for distributed estimation. *Signal Process.* **2016**, *122*, 93–97. [[CrossRef](#)]
11. Zheng, Z.; Liu, Z.; Huang, M. Diffusion least mean square/fourth algorithm for distributed estimation. *Signal Process.* **2017**, *134*, 268–274. [[CrossRef](#)]
12. Ni, J.; Chen, J.; Chen, X. Diffusion sign-error LMS algorithm: Formulation and stochastic behavior analysis. *Signal Process.* **2016**, *128*, 142–149. [[CrossRef](#)]
13. Liu, W.; Pokharel, P.P.; Principe, J.C. Correntropy: Properties and Applications in Non-Gaussian Signal Processing. *IEEE Trans. Signal Process.* **2007**, *55*, 5286–5298. [[CrossRef](#)]
14. Chen, B.; Principe, J.C. Maximum Correntropy Estimation Is a Smoothed MAP Estimation. *IEEE Signal Process. Lett.* **2012**, *19*, 491–494. [[CrossRef](#)]

15. Ma, W.; Chen, B.; Duan, J.; Zhao, H. Diffusion maximum correntropy criterion algorithms for robust distributed estimation. *Digital Signal Process.* **2016**, *58*, 10–19. [[CrossRef](#)]
16. Guo, Y.; Ma, B.; Li, Y. A Kernel-Width Adaption Diffusion Maximum Correntropy Algorithm. *IEEE Access* **2020**, *8*, 33574–33587. [[CrossRef](#)]
17. Guo, Y.; Li, J.; Li, Y. Diffusion Correntropy Subband Adaptive Filtering (SAF) Algorithm over Distributed Smart Dust Networks. *Symmetry* **2019**, *11*, 1335. [[CrossRef](#)]
18. Song, P.; Zhao, H.; Li, P.; Shi, L. Diffusion affine projection maximum correntropy criterion algorithm and its performance analysis. *Singal Process.* **2021**, *181*, 1078918. [[CrossRef](#)]
19. Chen, B.; Xing, L.; Zhao, H.; Zheng, N.; Principe, J.C. Generalized correntropy for robust adaptive filtering. *IEEE Trans. Signal Process.* **2016**, *64*, 3376–3387. [[CrossRef](#)]
20. Zhao, J.; Zhang, H. Kernel Recursive Generalized Maximum Correntropy. *IEEE Signal Process. Lett.* **2017**, *24*, 1832–1836. [[CrossRef](#)]
21. Hakimi, S.; Hodtani, G.A. Generalized maximum correntropy detector for non-Gaussian environments. *Int. J. Adapt. Control Signal Process.* **2018**, *32*, 83–97. [[CrossRef](#)]
22. Zhang, T.; Wan, S. Nystrom Kernel under Generalized Maximum Correntropy Criterion. *IEEE Signal Process. Lett.* **2020**, *27*, 1535–1539. [[CrossRef](#)]
23. Luo, X.; Sun, J.; Wang, L.; Wang, W.; Zhao, W.; Wu, J.; Wang, J.-H.; Zhang, Z. Short-Term Wind Speed Forecasting via Stacked Extreme Learning Machine with Generalized Correntropy. *IEEE Trans. Indust. Inform.* **2018**, *14*, 4963–4971. [[CrossRef](#)]
24. Ma, W.; Qiu, J.; Liu, X.; Xiao, G.; Duan, J.; Chen, B. Unscented Kalman Filter with Generalized Correntropy Loss for Robust Power System Forecasting-Aided State Estimation. *IEEE Trans. Indust. Inform.* **2019**, *15*, 6091–6100. [[CrossRef](#)]
25. Chen, F.; Li, X.; Wang, L.; Wu, J. Diffusion generalized maximum correntropy criterion algorithm for distributed estimation over multitask network. *Digital Signal Process.* **2018**, *81*, 16–25. [[CrossRef](#)]
26. Mial, D.; Jia, L.; Zhang, J. Robust diffusion affine projection generalized maximum correntropy criterion algorithm over networks. In Proceedings of the 2021 40th Chinese Control Conference, Shanghai, China, 26–28 July 2021.
27. Chen, B.; Wang, X.; Li, Y.; Principe, J.C. Maximum Correntropy Criterion with Variable Center. *IEEE Signal Process. Lett.* **2019**, *26*, 1212–1216. [[CrossRef](#)]
28. Zhu, L.; Song, C.; Pan, L.; Li, J. Adaptive filtering under the maximum correntropy criterion with variable center. *IEEE Access.* **2019**, *7*, 105902–105908. [[CrossRef](#)]
29. Liu, X.; Song, C.; Pang, Z. Kernel recursive maximum correntropy with variable center. *Signal Process.* **2021**, *191*, 108364. [[CrossRef](#)]
30. Yang, J.; Cao, J.; Xue, A. Robust Maximum Mixture Correntropy Criterion-Based Semi-Supervised ELM with Variable Center. *IEEE Trans. Circuits Syst. II Express Briefs* **2020**, *67*, 3572–3576. [[CrossRef](#)]
31. Sun, Q.; Zhang, H.; Wang, X.; Ma, W.; Chen, B. Sparsity Constrained Recursive Generalized Maximum Correntropy Criterion with Variable Center Algorithm. *IEEE Trans. Circuits Syst. II Express Briefs* **2020**, *67*, 3517–3521. [[CrossRef](#)]