

Article

Diffusion Correntropy Subband Adaptive Filtering (SAF) Algorithm over Distributed Smart Dust Networks

Ying Guo ¹, Jingjing Li ¹ and Yingsong Li ^{2,3,*} 

¹ School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China; gy20072009@163.com (Y.G.); ljj02062029@163.com (J.L.)

² College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

³ Key Laboratory of Microwave Remote Sensing, Chinese Academy of Sciences, Beijing 100190, China

* Correspondence: liyingsong@ieee.org

Received: 6 September 2019; Accepted: 18 October 2019; Published: 26 October 2019



Abstract: The diffusion subband adaptive filtering (DSAF) algorithm has attracted much attention in recent years due to its decorrelation ability for colored input signals. In this paper, a modified DSAF algorithm using the symmetry maximum correntropy criterion (MCC) with individual weighting factors is proposed and discussed to combat impulsive noise, which is denoted as the MCC-DSAF algorithm. During the iterations, the negative exponent in the Gaussian kernel of the MCC-DSAF eliminates the interference of outliers to provide a robust performance in non-Gaussian noise environments. Moreover, in order to enhance the convergence for sparse system identifications, a variant of MCC-DSAF named as improved proportionate MCC-DSAF (MCC-IPDSAF) is presented and investigated, which provides a dynamic gain assignment matrix in the MCC-DSAF to adjust the weighted values of each coefficient. Simulation results verify that the newly presented MCC-DSAF and MCC-IPDSAF algorithms are superior to the popular DSAF algorithms.

Keywords: diffusion subband adaptive filtering algorithm; maximum correntropy criterion; impulsive noise; proportionate adaptive filters ; sparse system identifications

1. Introduction

Distributed estimation is required to get the interested estimated parameters from the data collected in distributed networks and sensors [1], and has been widely investigated and utilized in wireless sensor networks, target locations, environmental monitorings, medical applications, military applications, and other fields [2–5]. Diffusion strategy exchanges information between a current node and its neighbors in the networks, which makes it widely studied in distributed estimation [6]. A global solution in the network performs iterative operations when all node information is fused to the center, which requires a large amount of energy and communication resources [7]. Thus, many adaptive filtering (AF) algorithms in light of distributed estimation were reported like diffusion affine projection algorithms (DAPA) [8], diffusion least mean square algorithms (DLMS) [9], and diffusion subband adaptive filtering (SAF) algorithms (DSAF) [10]. DLMS was developed in line with the least mean square (LMS) for distributed estimation since the LMS is simple, but its convergence speed will sharply deteriorate for colored input signals. Although the DAPA and DSAF can improve the convergence when the input is colored, the complexity for DAPA is increased. DSAF is more popular because of its simple computational complexity that is similar to the DLMS [11–15]. However, the above mentioned algorithms are all developed based on the l_2 -norm optimization criterion, whose performance will be seriously degraded under non-Gaussian interference.

Recently, the l_1 norm optimization criterion (LOC) was used for dealing with strong impulsive noise to provide good anti-interference abilities [16]. Then, AF algorithms in line of the LOC were reported, such as the diffusion sign subband AF algorithm (DSSAF) [17], the DSSAF with enlarged cooperation (DSSAF-EC) [18], and the individual weighting factored DSSAF algorithm (IWF-DSSAF) [19]. The above mentioned algorithms are very sensitive to non-Gaussian noise, resulting in potential deterioration of their steady-state error performance, since the LOC algorithms use the sign function on the error in each iterative process. In order to ensure the reliability of signal estimation and transmission in real-life systems, it is necessary to construct a more robust method to suppress non-Gaussian noise interference.

Since the maximum correntropy scheme can resist the outliers, it has been used to deal with non-Gaussian noise in recent years [20]. For example, the diffusion maximum correntropy criterion (DMCC) [21–25] and the diffusion minimum error entropy (DMEE) algorithms [22] provide good robustness under non-Gaussian noise-interference, and have been developed over networks based on the correntropy theory. Their performance demonstrates that the computational complexity of symmetry maximum correntropy criterion (MCC)-based algorithms is less than that of minimum error entropy (MEE)-based algorithms.

Moreover, the impulse responses (IR) of many scenes are regarded as sparse [26–31], which means that the values of most IR coefficients are zero or very small, where only a small part of the IR coefficients have significant amplitude. Then, proportionate techniques [32,33] have been used to enhance the DSAF algorithms for sparse system identifications. The main idea of the proportionate adaption AF algorithms is that the proportionate assignment matrix assigns different gain to each coefficient. The large coefficients obtain large gain to provide large step-size, and vice versa. The proportionate-type scheme has been utilized in diffusion proportionate sign subband adaptive filtering (DPSSAF) [34] and its series of variants [18,19]. In this paper, the MCC [23,35,36] is adopted for constructing a new cost function in this paper, where the proposed novel algorithm is named as MCC-DSAF. The outliers in the noise make error become large, and the negative exponential term of the Gaussian kernel function in the MCC-DSAF causes the outlier to approach zero. Therefore, compared with DSSAFs, MCC-DSAF has good adaptability and stronger robustness in non-Gaussian noise environments. We also proposed an improved proportionate MCC-DSAF (MCC-IPDSAF) by giving each coefficient an independent gain. The MCC-IPDSAF not only has the ability to resist impulsive noise, but also is suitable for system identifications with different sparse characteristics. The simulated results show that the derived two algorithms have good performance if the convergence performance, steady state error and tracking behaviors are compared with popular methods.

2. The Proposed Algorithms

A smart dust network is a system that contains many sensors, robots, or other devices to build a tiny microelectromechanical systems (MEMS) to detect light, temperature, vibration, magnetism, or chemicals. Herein, an N -node smart dust network within the wireless sensor network framework is considered. At node n and time slot l , the signal $d_n(l)$ arises from the linear model:

$$d_n(l) = \mathbf{u}_n^T(l) \mathbf{w}_0 + \eta_n(l), n = 1, 2, 3, 4, \dots, N \quad (1)$$

where $\mathbf{u}_n(l) = [u_n(l), u_n(l-1), \dots, u_n(l-M+1)]^T$ denotes input vector, \mathbf{w}_0 is the unknown vector with length M that we expect to estimate, and $\eta_n(l)$ is the system noise.

2.1. Review of the DSAF Algorithm

The DSAF is implemented based on a multiband structure, which is shown Figure 1. In Figure 1, $\uparrow I$ and $\downarrow I$ represent N interpolations and decimations, respectively. Considering $H_i(z) = \sum_{m=0}^{M-1} h_i(m)z^{-m}$, the desired signal $d_n(l)$ and inputting signal $u_n(l)$ for node n are assigned to I subband

signals $d_{n,i}(l)$ and $u_{n,i}(l)$, where $i = 0, 1, \dots, I - 1$, I is the number of subbands. The subband desired vectors and subband input vectors for node n are expressed as:

$$\mathbf{u}_{n,i}(k) = [u_{n,i}(kI), u_{n,i}(kI - 1), \dots, u_{n,i}(kI - M + 1)]^T, \tag{2}$$

$$\mathbf{d}_n(k) = [d_{n,0,D}(k), d_{n,1,D}(k), \dots, d_{n,I-1,D}(k)]^T. \tag{3}$$

$u_{n,i}(l)$ for node n are handled by adaptive filter $\hat{W}_n(z)$ whose weight vector is $w_n(l)$, which is to generate output signal $y_{n,i}(l)$. $d_{n,i}(l)$ for node n and the output signals $y_{n,i}(l)$ are to get $d_{n,i,D}(k)$ and $y_{n,i,D}(k)$. The subband errors $e_{n,i,D}(k)$ are the difference between $d_{n,i,D}(k)$ and $y_{n,i,D}(k)$. Using the synthesis filter bank $G_i(z) = \sum_{m=0}^{M-1} g_i(m)z^{-m}$, the fullband error $e_n(l)$ is obtained. The original and decimated sequences are denoted by variables l and k respectively, where $l = kI$.

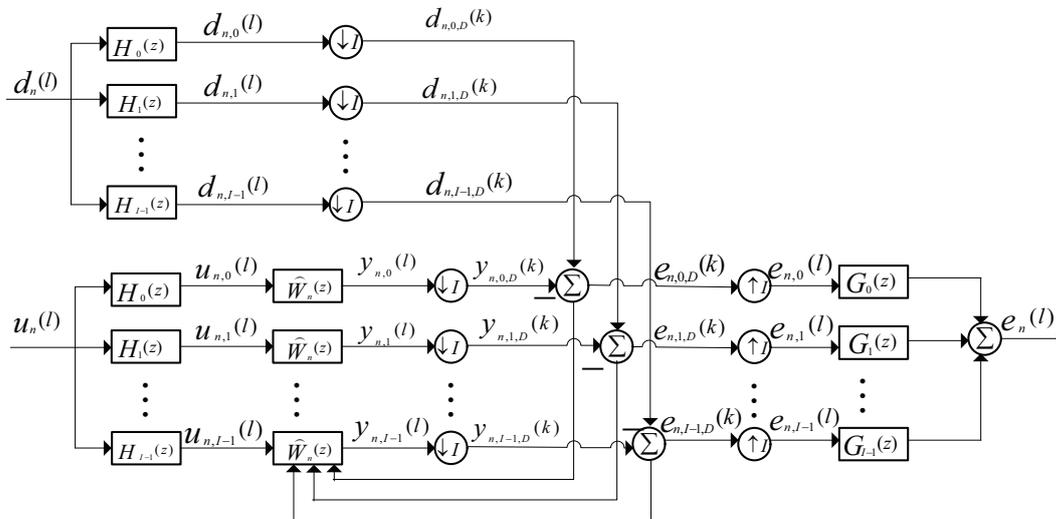


Figure 1. The scheme of the diffusion subband adaptive filtering (DSAF) algorithm.

As we know, DSAF algorithms can be divided into two types, namely (Adapt-then-Combine) ATC and (Combine-then-Adapt) CTA [37]. In general, ATC performs better than CTA, and hence, the ATC type is employed herein.

For DSAF, its cost function for node n is

$$J(k) = E[\|\sum_{i=0}^{L-1} (\mathbf{d}_{n,i,D}(k) - \mathbf{u}_{n,i}^T(k)\mathbf{w}_n(k-1))\|^2], \tag{4}$$

in which E denotes the expected symbol. By using ATC type in Equation (4), $J(k)$ is minimized to obtain the updated equation of the DSAF:

$$\begin{cases} \boldsymbol{\phi}_n(k) = \mathbf{w}_n(k-1) + 2\mu_n \sum_{i=0}^{L-1} \frac{\mathbf{u}_{n,i}(k)e_{n,i,D}(k)}{\mathbf{u}_{n,i}^T(k)\mathbf{u}_{n,i}(k) + \varepsilon} \\ \mathbf{w}_n(k) = \sum_{j \in \mathcal{N}_n} c_{jn}\boldsymbol{\phi}_j(k) \end{cases}, \tag{5}$$

where $e_{n,i,D}(k)$ is

$$e_{n,i,D}(k) = d_{n,i,D}(k) - \mathbf{u}_{n,i}^T(k)\mathbf{w}_n(k-1). \tag{6}$$

\mathcal{N}_n represents the neighbor of node n . c_{jn} denotes the combination coefficients of the $N \times N$ matrix \mathbf{C} , where, $\sum_{j \in \mathcal{N}_n} c_{jn} = 1$ and $c_{jn} = 0$ if $j \notin \mathcal{N}_n$. In other words, c_{jn} is zero when j is not connected to n . Otherwise, the column and row of the combination coefficients c_{jn} in \mathbf{C} are added to one. ε is a

regularization parameter which prevents the denominator from being zero, μ_n is the step-size for node n . $\phi_n(k)$ and $w_n(k)$ are the intermediate estimation and estimation of w_0 for node n .

Under the Gaussian noise interference, the DSAF has a decorrelation ability for the colored input signals and enjoys a faster convergence speed than DLMS. However, for the interference of non-Gaussian noise, the error sequences in the DSAF are not stabilized due to the pulse characteristics of non-Gaussian noise, which affects the convergence characteristics of DSAF. Then, a modified DSAF algorithm based on an MCC scheme with an individual weighting factor for improving the DSAF is proposed and given the name MCC-DSAF.

2.2. The Proposed MCC-DSAF Algorithm

Correntropy between two random variables X and Y is defined as

$$C(X, Y) = E[\kappa_\sigma(X - Y)] = \iint \kappa(x - y) f_{X,Y}(x, y) dx dy, \quad (7)$$

where κ_σ denotes the Mercer kernel, and $f_{X,Y}(x, y)$ is the joint probability density function. A Gaussian kernel is always used and is presented as

$$\kappa_\sigma(x - y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - y)^2}{2\sigma^2}\right), \quad (8)$$

where σ is the kernel size. To use the correntropy in our algorithm, we define $e_{n,i,D}(k) = x - y$, $x = d_{n,i,D}(k)$ and $y = \mathbf{u}_{n,i}^T(k) \mathbf{w}_n(k - 1)$. Then, the new cost function is defined as:

$$J^{glob}(k) = \sum_{n=1}^N \sum_{i=0}^{L-1} \left(\sqrt{\frac{\beta}{\pi}} \exp\left[-\beta \frac{(d_{n,i,D}(k) - \mathbf{u}_{n,i}^T(k) \mathbf{w}_n(k - 1))^2}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1}\right] \right), \quad (9)$$

where β denotes the kernel parameter related to the kernel size and $\beta = \frac{1}{2\sigma^2}$, $\varepsilon_1 > 0$ has a small value.

In the global network, a large amount of communication resources and energy are required and the real-time requirements of the system are high. Thus, the above problem should be well solved in the distributed network, and the global cost function should be changed into a local cost function.

For diffusion networks, local cost function can be formulated as linear-combination of local weighted correntropy, which is expressed as:

$$J_n^{loc}(k) = \sum_{j \in \mathcal{N}_n} c_{jn} \left(\sum_{i=0}^{L-1} \left(\sqrt{\frac{\beta}{\pi}} \exp\left[-\beta \frac{(d_{n,i,D}(k) - \mathbf{u}_{n,i}^T(k) \mathbf{w}_n(k - 1))^2}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1}\right] \right) \right). \quad (10)$$

Driven by Equation (10), the increment of weight vector $\Delta \mathbf{w}_n(k)$ at instant time k is written as:

$$\begin{aligned} \Delta \mathbf{w}_n(k) &= \frac{\partial J_n^{loc}(k)}{\partial \mathbf{w}} \\ &= \sum_{j \in \mathcal{N}_n} c_{jn} \left(\sum_{i=0}^{L-1} \left(2\sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{(d_{n,i,D}(k) - \mathbf{u}_{n,i}^T(k) \mathbf{w}_n(k - 1))^2}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1}\right] \right) \right. \\ &\quad \times \left. \frac{\mathbf{u}_{n,i}(k) (d_{n,i,D}(k) - \mathbf{u}_{n,i}^T(k) \mathbf{w}_n(k - 1))}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1} \right) \\ &= \sum_{j \in \mathcal{N}_n} c_{jn} \left(\sum_{i=0}^{L-1} \left(2\sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1}\right] \right) \right. \\ &\quad \times \left. \frac{\mathbf{u}_{n,i}(k) e_{n,i,D}(k)}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1} \right). \end{aligned} \quad (11)$$

Then, the updating of the MCC-DSAF based on the gradient method can be obtained,

$$\begin{aligned} \mathbf{w}_n(k) &= \mathbf{w}_n(k-1) + \mu_n \Delta \mathbf{w}_n(k) \\ &= \mathbf{w}_n(k-1) + \mu_n \sum_{j \in \mathcal{N}_n} c_{jn} \left(\sum_{i=0}^{L-1} \left(2 \sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1}\right] \right) \right. \\ &\quad \left. \times \frac{\mathbf{u}_{n,i}(k) e_{n,i,D}(k)}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1} \right). \end{aligned} \quad (12)$$

For the diffusion strategy, the local estimation $\mathbf{w}_n(k)$ is a linear combination of intermediate-estimation $\boldsymbol{\phi}_n(k)$, their relationship can be expressed

$$\mathbf{w}_n(k) = \sum_{j \in \mathcal{N}_n} c_{jn} \boldsymbol{\phi}_n(k). \quad (13)$$

In such way, $\Delta \mathbf{w}_n(k)$ is written as:

$$\Delta \mathbf{w}_n(k) = \sum_{j \in \mathcal{N}_n} c_{jn} \Delta \boldsymbol{\phi}_n(k). \quad (14)$$

Comparing Equation (11) with Equation (14), the increment of intermediate estimation $\Delta \boldsymbol{\phi}_n(k)$ is

$$\Delta \boldsymbol{\phi}_n(k) = 2 \sqrt{\frac{\beta^3}{\pi}} \sum_{i=0}^{L-1} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1}\right] \times \frac{\mathbf{u}_{n,i}(k) e_{n,i,D}(k)}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1}. \quad (15)$$

From Equation (12), the updating of the intermediate-estimation $\boldsymbol{\phi}_n(k)$ is

$$\boldsymbol{\phi}_n(k) = \boldsymbol{\phi}_n(k-1) + \mu_n \Delta \boldsymbol{\phi}_n(k). \quad (16)$$

From Equation (13), $\mathbf{w}_n(k-1)$ contains more data information from neighbor nodes compared with $\boldsymbol{\phi}_n(k-1)$ [7,22]. According to the diffusion in [7], by replacing $\boldsymbol{\phi}_n(k-1)$ by $\mathbf{w}_n(k-1)$ in Equation (16), we have

$$\boldsymbol{\phi}_n(k) = \mathbf{w}_n(k-1) + \mu_n \Delta \boldsymbol{\phi}_n(k). \quad (17)$$

Using Equations (12), (15), and (17), the updating equation of the MCC-DSAF is rewritten as:

$$\left\{ \begin{aligned} \boldsymbol{\phi}_n(k) &= \mathbf{w}_n(k-1) + 2\mu_n \sqrt{\frac{\beta^3}{\pi}} \sum_{i=0}^{L-1} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1}\right] \\ &\quad \times \frac{\mathbf{u}_{n,i}(k) e_{n,i,D}(k)}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1} \\ \mathbf{w}_n(k) &= \sum_{j \in \mathcal{N}_n} c_{jn} \boldsymbol{\phi}_j(k). \end{aligned} \right. \quad (18)$$

When the error $e_{n,i,D}(k)$ is interrupted by the pulse during the weight updating process, the negative exponential action in the Gaussian kernel function term $\exp\left[-\beta \frac{e_{n,i,D}^2(k)}{\mathbf{u}_{n,i}^T(k) \mathbf{u}_{n,i}(k) + \varepsilon_1}\right]$ makes the outlier close to zero, which ensures that the MCC-DSAF has good performance under non-Gaussian noise interference.

2.3. The Proposed MCC-IPDSAF Algorithm

The MCC-IPDSAF is proposed using the adaptive gain matrix in the MCC-DSAF to obtain faster convergence when the unknown system is sparse since the adaptive gain matrix $G_n(k)$ can get a better balance between the convergence and steady state error. The update equation for the MCC-IPDSAF is

$$\begin{cases} \phi_n(k) = w_n(k-1) + 2\mu_n \sqrt{\frac{\beta^3}{\pi}} \sum_{i=0}^{L-1} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{\mathbf{u}_{n,i}^T(k)\mathbf{u}_{n,i}(k) + \varepsilon_1}\right] \\ \quad \times \frac{\mathbf{G}_n(k)\mathbf{u}_{n,i}(k)e_{n,i,D}(k)}{\mathbf{u}_{n,i}^T(k)\mathbf{G}_n(k)\mathbf{u}_{n,i}(k) + \varepsilon_1} \\ w_n(k) = \sum_{j \in \mathcal{N}_n} c_{jn} \phi_j(k) \end{cases} \quad (19)$$

There are several methods for choosing $G_n(k)$ to render it suitable for a sparse system [33]. As we know, the main idea of adaptive gain matrix in AFs is to obtain larger step sizes for the larger filter coefficients, which guarantees their convergence. Therefore, the MCC-IPDSAF can provide a quicker convergence compared to the MCC-DSAF in the sparse system.

$G_n(k) = \text{diag}[g_{n,0}(k), g_{n,1}(k), \dots, g_{n,M-1}(k)]$ denotes the $M \times M$ adaptive gain matrix. The diagonal elements of $G_n(k)$ are calculated by

$$g_{n,m}(k) = \frac{1 - \alpha}{2M} + (1 + \alpha) \frac{|w_{n,m}(k)|}{\|\mathbf{w}_n(k)\|_1 + \varepsilon_2}, \quad (20)$$

where α is a parameter related to system sparsity. ε_2 is a regularization to prevent the denominator from being zero.

3. Performance Analysis

3.1. Data Model and Assumption

Mean square analysis of the MCC-DSAF will be performed. We define the following global variables:

$$\mathbf{w}(k) = \text{col}\{\mathbf{w}_1(k), \mathbf{w}_2(k), \dots, \mathbf{w}_N(k)\}, \quad (21)$$

$$\boldsymbol{\phi}(k) = \text{col}\{\boldsymbol{\phi}_1(k), \boldsymbol{\phi}_2(k), \dots, \boldsymbol{\phi}_N(k)\}, \quad (22)$$

$$\mathbf{U}(k) = \text{diag}\{\mathbf{u}_1(k), \mathbf{u}_2(k), \dots, \mathbf{u}_N(k)\}, \quad (23)$$

$$\mathbf{u}_n(k) = \text{col}\{\mathbf{u}_{n,0}(k), \mathbf{u}_{n,1}(k), \dots, \mathbf{u}_{n,L-1}(k)\}, \quad (24)$$

$$\mathbf{d}(k) = \text{col}\{d_{1,0}(k), \dots, d_{1,L-1}(k), d_{N,0}(k), \dots, d_{N,L-1}(k)\}, \quad (25)$$

$$\mathbf{v}(k) = \text{col}\{\eta_{1,0}(k), \dots, \eta_{1,L-1}(k), \eta_{N,0}(k), \dots, \eta_{N,L-1}(k)\}, \quad (26)$$

where $\text{col}\{\cdot\}$ denotes the column vector, and $\text{diag}\{\cdot\}$ denotes a diagonal matrix. The desired signal of the entire network $\mathbf{d}(k) = \mathbf{U}^T(k)\mathbf{w}^0 + \mathbf{v}(k)$, and $\mathbf{w}^0 = \mathbf{I}\mathbf{w}_0$, $\mathbf{I} = \text{col}\{\mathbf{I}_M, \mathbf{I}_M, \dots, \mathbf{I}_M\}$ is an $MN \times M$ matrix.

$\mathbf{H}(k)$ is

$$\begin{cases} \mathbf{H}(k) = \text{diag}\left\{\frac{1}{\mathbf{u}_{1,0}^T(k)\mathbf{u}_{1,0}(k) + \varepsilon_1}, \dots, \frac{1}{\mathbf{u}_{1,L-1}^T(k)\mathbf{u}_{1,L-1}(k) + \varepsilon_1}, \right. \\ \left. \dots, \frac{1}{\mathbf{u}_{N,0}^T(k)\mathbf{u}_{N,0}(k) + \varepsilon_1}, \dots, \frac{1}{\mathbf{u}_{N,L-1}^T(k)\mathbf{u}_{N,L-1}(k) + \varepsilon_1}\right\}. \end{cases} \quad (27)$$

Herein, the \mathbf{P} matrix is a collection of local step parameters:

$$\mathbf{P} = \text{diag}\{\mu_1, \mu_2, \dots, \mu_N\} \otimes \mathbf{I}_M, \quad (28)$$

where \otimes denotes the Kronecker product. Next, \mathbf{C} is the combination coefficients matrix of $N \times N$:

$$\mathbf{C} = \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,N} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,N} \\ \vdots & \vdots & \cdots & \vdots \\ c_{N,1} & c_{N,2} & \cdots & c_{N,N} \end{bmatrix}, \quad (29)$$

and the matrix \mathbf{A} is defined as:

$$\mathbf{A} = \mathbf{C}^T \otimes \mathbf{I}_M. \quad (30)$$

The matrix $\mathbf{\Omega}(k)$ is given by

$$\mathbf{\Omega}(k) = \begin{Bmatrix} 2\sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{e_{1,0,D}^2(k)}{\mathbf{u}_{1,0}^T(k)\mathbf{u}_{1,0}(k) + \varepsilon_1}\right], \dots, \\ 2\sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{e_{1,L-1,D}^2(k)}{\mathbf{u}_{1,L-1}^T(k)\mathbf{u}_{1,L-1}(k) + \varepsilon_1}\right], \dots, \\ 2\sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{e_{N,0,D}^2(k)}{\mathbf{u}_{N,0}^T(k)\mathbf{u}_{N,0}(k) + \varepsilon_1}\right], \dots, \\ 2\sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{e_{N,L-1,D}^2(k)}{\mathbf{u}_{N,L-1}^T(k)\mathbf{u}_{N,L-1}(k) + \varepsilon_1}\right] \end{Bmatrix}. \quad (31)$$

Using the global variables, the global update equation of the MCC-DSAF can be obtained,

$$\begin{cases} \boldsymbol{\phi}(k) = \mathbf{w}(k-1) + \mathbf{P}\mathbf{\Omega}(k)\mathbf{H}(k)\mathbf{U}(k)[\mathbf{d}(k) - \mathbf{U}^T(k)\mathbf{w}(k-1)] \\ \mathbf{w}(k) = \mathbf{A}\boldsymbol{\phi}(k). \end{cases} \quad (32)$$

In the analysis, we have the following assumptions:

Assumption 1. All input regressions $\mathbf{u}_n(l)$ are independent. $\eta_n(l)$ is independent and independent of $\mathbf{u}_n(l)$.

Assumption 2. Subband colored input signal $\mathbf{u}_{n,i}(k)$ is close to a white signal.

Assumption 3. The Gaussian kernel function $\left\{2\sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{\mathbf{u}_{n,i}^T(k)\mathbf{u}_{n,i}(k) + \varepsilon_1}\right]\right\}$ is independent of $\mathbf{u}_{n,i}(k)$.

Assumption 1 is widely used in the analysis of AF algorithms. Assumption 3 does not really apply to the proposed algorithms, since $\left\{2\sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{\mathbf{u}_{n,i}^T(k)\mathbf{u}_{n,i}(k) + \varepsilon_1}\right]\right\}$ is an error function. In [21,37–39], the variable step size is independent of $\mathbf{u}_{n,i}(k)$. $\left\{2\sqrt{\frac{\beta^3}{\pi}} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{\mathbf{u}_{n,i}^T(k)\mathbf{u}_{n,i}(k) + \varepsilon_1}\right]\right\}$ is considered as a variable step term.

The weighted error-vector for node n is expressed as:

$$\tilde{\mathbf{w}}_n(k) = \mathbf{w}_0 - \mathbf{w}_n(k). \quad (33)$$

The global weighted error-vector is:

$$\tilde{w}(k) = w^0 - w(k). \tag{34}$$

Due to $A = C^T \otimes I_M$ and $\|C\|_2 = 1$, where C is the matrix whose coefficients are c_{jn} , and we have $A = I_{MN}$. Thus, we get the relationship of $w^0 = Aw^0$. Replacing $w(k)$ with $A\phi(k)$ in Equation (32), then Equation (34) is modified to be:

$$\begin{aligned} \tilde{w}(k) &= w^0 - w(k) \\ &= Aw^0 - A\phi(k) \\ &= Aw^0 - A[w(k-1) + P\Omega(k)H(k)U(k)(d(k) - U^T(k)w(k-1))] \\ &= A\tilde{w}(k-1) - A[P\Omega(k)H(k)U(k)(U^T(k)w_0 + v(k) - U^T(k)w(k-1))] \\ &= A\tilde{w}(k-1) - A[P\Omega(k)H(k)U(k)(U^T(k)\tilde{w}(k-1) + v(k))] \\ &= A[I_{MN} - P\Omega(k)H(k)U(k)U^T(k)]\tilde{w}(k-1) - AP\Omega(k)H(k)U^T(k)v(k). \end{aligned} \tag{35}$$

Next, the mean square analysis for $\tilde{w}(k)$ will be presented.

3.2. Convergence Analysis

The expectation is simultaneously exerted on both sides of (35), then we get

$$\begin{aligned} E[\tilde{w}(k)] &= A[I_{MN} - E[P\Omega(k)H(k)U(k)U^T(k)]]E[\tilde{w}(k-1)] \\ &\quad - AE[P\Omega(k)H(k)U^T(k)v(k)]. \end{aligned} \tag{36}$$

It can be seen from Assumption 3 that $\Omega(k)$ is independent of $U(k)$. Thereby, we obtain

$$E[P\Omega(k)H(k)U(k)U^T(k)] = PE[\Omega(k)]R_U, \tag{37}$$

where $R_U = E[H(k)U(k)U^T(k)] = E[\sum_{i=0}^{L-1} \frac{u_{n,i}(k)u_{n,i}^T(k)}{u_{n,i}^T(k)u_{n,i}(k) + \epsilon_1}]$.

From Assumption 1, we can find that the expectation of the last term of (36) is close to zero. Therefore, (36) is rewritten as

$$E[\tilde{w}(k)] = A[I_{MN} - PE[\Omega(k)]R_U]E[\tilde{w}(k-1)]. \tag{38}$$

In Equation (38), the matrix $\{A[I_{MN} - PE[\Omega(k)]R_U]\}$ should be stable. Thus, $\left\{ I - \mu_n E \left[2\sqrt{\frac{\beta^3}{\pi}} \sum_{i=0}^{L-1} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{u_{n,i}^T(k)u_{n,i}(k) + \epsilon_1}\right] \right] R_U \right\}$ should be stable for all n , which means that μ_n should satisfy the following equation

$$-1 < 1 - \mu_n \lambda_{\max}(R_U) E \left[2\sqrt{\frac{\beta^3}{\pi}} \sum_{i=0}^{L-1} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{u_{n,i}^T(k)u_{n,i}(k) + \epsilon_1}\right] \right] < 1, \tag{39}$$

then,

$$\begin{aligned} 0 < \mu_n < \frac{2}{\lambda_{\max}(R_U) E \left[2\sqrt{\frac{\beta^3}{\pi}} \sum_{i=0}^{L-1} \exp\left[-\beta \frac{e_{n,i,D}^2(k)}{u_{n,i}^T(k)u_{n,i}(k) + \epsilon_1}\right] \right]} \\ n = 1, 2, \dots, N, \end{aligned} \tag{40}$$

where λ_{\max} denotes the maximum eigenvalue of \mathbf{R}_U . If the l_1 norm of the weight $\mathbf{w}_n(k)$ is smaller than τ , we have

$$\begin{aligned} |e_{n,i,D}(k)| &= |d_{n,i,D} - \mathbf{u}_{n,i}^T(k)\mathbf{w}_n(k-1)| \\ &\leq \|\mathbf{u}_{n,i}(k)\|_1 \|\mathbf{w}_n(k-1)\|_1 + |d_{n,i,D}(k)| \\ &\leq \tau \|\mathbf{u}_{n,i}(k)\|_1 + |d_{n,i,D}(k)|. \end{aligned} \tag{41}$$

Next, let $\mathbf{u}_{n,i}^T(k)\mathbf{u}_{n,i}(k) = \|\mathbf{u}_{n,i}(k)\|_2^2$. Therefore, the following condition for the stability of the MCC-DSAF should satisfy

$$0 < \mu_n < \frac{2}{\lambda_{\max}(\mathbf{R}_U) E\left[2\sqrt{\frac{\beta^3}{\pi}} \sum_{i=0}^{L-1} \exp\left[-\beta \frac{[\tau \|\mathbf{u}_{n,i}(k)\|_1 + |d_{n,i,D}(k)|]^2}{\|\mathbf{u}_{n,i}(k)\|_2^2 + \varepsilon_1}\right]\right]} \tag{42}$$

$n = 1, 2, \dots, N.$

3.3. Steady-State Performance

Steady-state behavior of MCC-DSAF is studied herein. Σ denotes any symmetric positive definite weighting matrix, $\|\mathbf{t}\|_{\Sigma}^2$ denotes the weighting squared Euclidean norm $\|\mathbf{t}\|_{\Sigma}^2 = \mathbf{t}^T \Sigma \mathbf{t}$. Then, considering the Σ -Euclidean-norm on both sides of Equation (35) results in

$$\|\tilde{\mathbf{w}}(k)\|_{\Sigma}^2 = \|\tilde{\mathbf{w}}(k-1)\|_{\Sigma}^2 + \mathbf{v}^T(k)\mathbf{Y}(k)\mathbf{v}(k), \tag{43}$$

where

$$\mathbf{Y}(k) = \mathbf{U}^T(k)\mathbf{H}^T(k)\mathbf{\Omega}^T(k)\mathbf{P}^T\mathbf{A}^T\Sigma\mathbf{A}\mathbf{P}\mathbf{\Omega}(k)\mathbf{H}(k)\mathbf{U}(k), \tag{44}$$

and

$$\Sigma' = \mathbf{A}^T\Sigma\mathbf{A} - \mathbf{A}^T\Sigma\mathbf{A}\mathbf{P}\mathbf{Z}(k) - \mathbf{Z}^T(k)\mathbf{P}^T\mathbf{A}^T\Sigma\mathbf{A} + \mathbf{Z}^T(k)\mathbf{P}^T\mathbf{A}^T\Sigma\mathbf{A}\mathbf{P}\mathbf{Z}(k), \tag{45}$$

and

$$\mathbf{Z}(k) = \mathbf{\Omega}(k)\mathbf{H}(k)\mathbf{U}(k)\mathbf{U}^T(k). \tag{46}$$

We take the expectation on both sides of Equation (43) and find:

$$E[\|\tilde{\mathbf{w}}(k)\|_{\Sigma}^2] = E[\|\tilde{\mathbf{w}}(k-1)\|_{\Sigma}^2] + E[\mathbf{v}^T(k)\mathbf{Y}(k)\mathbf{v}(k)]. \tag{47}$$

$E[\Sigma']$ is

$$E[\Sigma'] = \mathbf{A}^T\Sigma\mathbf{A} - \mathbf{A}^T\Sigma\mathbf{A}PE[Z(k)] - E[Z^T(k)]\mathbf{P}^T\mathbf{A}^T\Sigma\mathbf{A} + E[Z^T(k)\mathbf{P}^T\mathbf{A}^T\Sigma\mathbf{A}\mathbf{P}\mathbf{Z}(k)]. \tag{48}$$

Herein, let $E[\Sigma'] = \Sigma''$. The $bvec\{\}$ operator is to convert a block matrix into a single-column vector. The \odot operator denotes the block Kronecker product. From $bvec[\mathbf{Q}\Sigma\mathbf{P}^T] = (\mathbf{P} \odot \mathbf{Q})\xi$ in [40], $\xi = bvec[\Sigma]$ and applying the $bvec\{\}$ operator to each item on the right side of (48) yields

$$bvec[\mathbf{A}^T\Sigma\mathbf{A}] = (\mathbf{A}^T \odot \mathbf{A}^T)\xi, \tag{49}$$

$$bvec[\mathbf{A}^T\Sigma\mathbf{A}PE[Z(k)]] = (E[Z^T(k)] \odot \mathbf{I}_{MN})(\mathbf{P}^T \odot \mathbf{I}_{MN})(\mathbf{A}^T \odot \mathbf{A}^T)\xi, \tag{50}$$

$$bvec[E[Z^T(k)]\mathbf{P}^T\mathbf{A}^T\Sigma\mathbf{A}] = (\mathbf{I}_{MN} \odot E[Z^T(k)])(\mathbf{I}_{MN} \odot \mathbf{P}^T)(\mathbf{A}^T \odot \mathbf{A}^T)\xi, \tag{51}$$

$$bvec[E[Z^T\mathbf{P}^T\mathbf{A}^T\Sigma\mathbf{A}\mathbf{P}\mathbf{Z}(k)]] = (E[Z^T(k) \odot \mathbf{Z}^T(k)])(\mathbf{P}^T \odot \mathbf{P}^T)(\mathbf{A}^T \odot \mathbf{A}^T)\xi. \tag{52}$$

By applying the $bvec\{\}$ operator to right side of Equation (48) without ξ , the final expression is represented by matrix \mathbf{Q} :

$$\mathbf{Q} = [\mathbf{I}_{M^2N^2} - (E[\mathbf{Z}(k)] \odot \mathbf{I}_{MN})(\mathbf{P}^T \odot \mathbf{I}_{MN}) - (\mathbf{I}_{MN} \odot E[\mathbf{Z}^T(k)]) (\mathbf{I}_{MN} \odot \mathbf{P}^T) + (E[\mathbf{Z}^T(k) \odot \mathbf{Z}^T(k)])(\mathbf{P}^T \odot \mathbf{P}^T)](\mathbf{A}^T \odot \mathbf{A}^T), \tag{53}$$

where $bvec[\Sigma] = \xi$, and applying the $bvec\{\}$ operator to left side of Equation (48), we get $bvec[\Sigma''] = \xi'$. Finally, we have

$$\xi' = \mathbf{Q}\xi. \tag{54}$$

Also, applying $bvec\{\}$ operator to the second term on the right side of Equation (47) yields

$$bvec[E[\mathbf{v}^T(k)\mathbf{Y}(k)\mathbf{v}(k)]] = \mathbf{R}\xi. \tag{55}$$

Thus, \mathbf{R} is

$$\mathbf{R} = E[(\mathbf{v}^T(k) \odot \mathbf{v}^T(k))(\mathbf{U}^T(k) \odot \mathbf{U}^T(k))(\mathbf{H}^T(k) \odot \mathbf{H}^T(k)) \times (\mathbf{\Omega}^T(k) \odot \mathbf{\Omega}^T(k))(\mathbf{P}^T \odot \mathbf{P}^T)(\mathbf{A}^T \odot \mathbf{A}^T)]. \tag{56}$$

According to the above analysis, Equation (47) is obtained,

$$E[\|\tilde{\mathbf{w}}(k)\|_{\xi}^2] = E[\|\tilde{\mathbf{w}}(k-1)\|_{\mathbf{Q}\xi}^2] + \mathbf{R}\xi. \tag{57}$$

The mean square deviation (MSD) for node n is given by

$$MSD_n = E[\|\tilde{\mathbf{w}}_n(k)\|^2] = E[\|\mathbf{w}_0 - \mathbf{w}_n(k)\|^2]. \tag{58}$$

Here, m_n is defined as

$$m_n = vec[diag(b_{n,N}) \otimes \mathbf{I}_M], \tag{59}$$

where $b_{n,N}$ denotes the n -th column vector of \mathbf{I}_N , vec stacks the columns of its matrix into a column vector. Let $\xi = m_n$ in Equation (57). When k approaches infinity, the MSD for node n is

$$MSD_n = \mathbf{R}[\mathbf{I} - \mathbf{Q}]^{-1}m_n. \tag{60}$$

The MSD of the entire network is defined by the average of all node MSDs:

$$MSD^{network} = \frac{1}{N} \sum_{n=1}^N MSD_n. \tag{61}$$

4. Simulation

The effectiveness of the proposed algorithms is verified through experimental simulation. Figure 2 shows the network topology with 20 nodes. The location coordinates for the nodes in a squared area are $[0, 1.2] \times [0, 1.2]$.

The combination coefficients c_{jn} are obtained by Metropolis criterion [11,41], where $c_{jn} = \frac{1}{\max(\mathcal{N}_n, \mathcal{N}_j)}$, if $n \neq j$ and $c_{jn} = 1 - \sum_{j \neq n} c_{jn}$, if $n = j$. The unknown system has a length of $M = 128$

and the calculation of sparsity is $\zeta(\mathbf{w}_0) = \frac{M}{M - \sqrt{M}} \left(1 - \frac{\|\mathbf{w}_0\|_1}{\sqrt{M}\|\mathbf{w}_0\|_2} \right)$ [42,43] with $\zeta(\mathbf{w}_0) = 0.751$.

Figure 3 shows the non-Gaussian noise distribution of $P_r = 0.01$ and $P_r = 0.1$ for node $n = 1$. The normalized mean square deviation (NMSD), $NMSD = 1/N \sum_n 10 \log_{10} \frac{\|\mathbf{w}_n - \mathbf{w}_0\|^2}{\|\mathbf{w}_0\|^2}$ is used to evaluate the behaviors of the devised algorithms. The smaller the value of NMSD, the closer the estimated results are to the unknown system. In order to make it easier to compare the differences

between the different algorithms, the convergence lines are run in 30 independent simulations to get the average values.

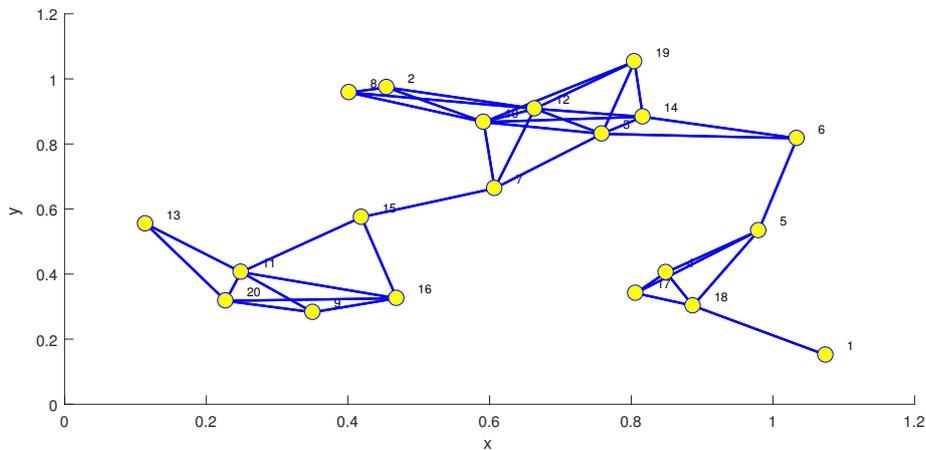


Figure 2. Diffusion network topology with 20 nodes within a squared area of $[0, 1.2] \times [0, 1.2]$.

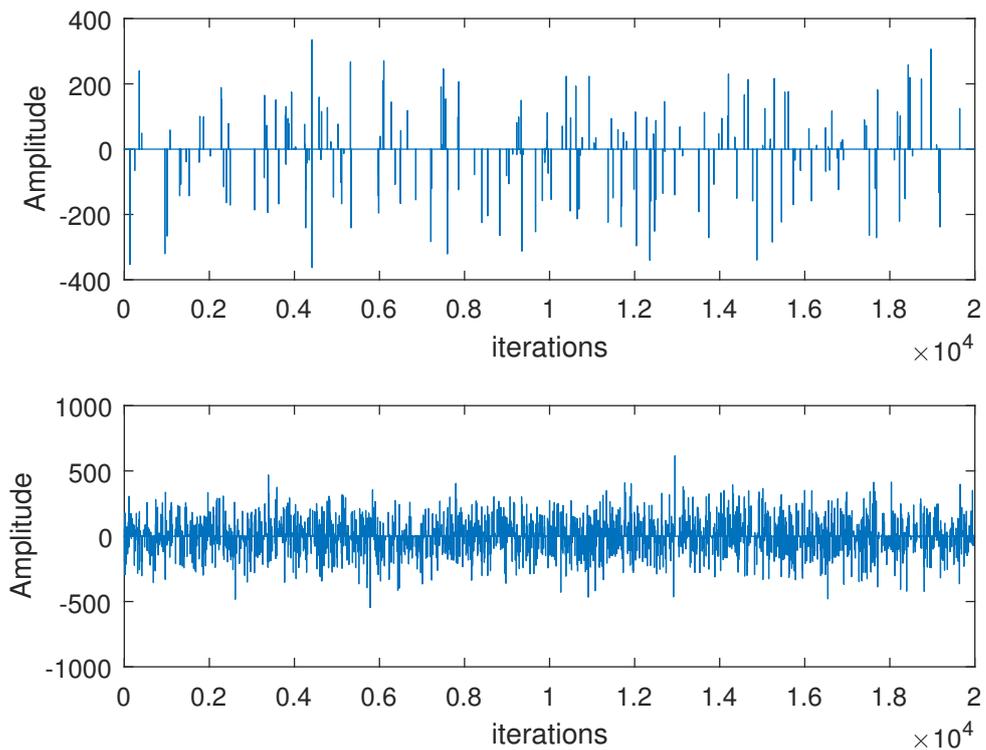


Figure 3. The non-Gaussian noise of $P_r = 0.01$ (up) and $P_r = 0.1$ (down) for node $n = 1$.

System Identification

The algorithms use four-subband cosine modulation filter banks. White, colored and speech signals are used as input in this section. The colored signal $u_n(l)$ is realized by Gaussian white noise via a first-order system with its transform function of $H(z) = \frac{1}{1 - 0.95z^{-1}}$. The variance of the input signal and the variance of the Gaussian noise are given in Figure 4. Zero-mean Gaussian noise $v_n(l)$ and the impulsive noise $z_n(l)$ are used to construct the measurement noise $\eta_n(l)$. The impulsive noise $z_n(l)$ is obtained by using the Bernoulli process $\phi_n(l)$ and the Gaussian process $q_n(l)$, which is defined

as $z_n(l) = q_n(l)\phi_n(l)$. The probability density function of the Bernoulli process is $P\{\phi_n = 0\} = 1 - P_r$, $P\{\phi_n = 1\} = P_r$. The signal-to-noise ratio (SNR) and signal-to-interference ratio (SIR) used in [18] are given by $SNR = 10\log_{10}\frac{\sigma_{d,n}^2}{\sigma_{v,n}^2}$ and $SIR = 10\log_{10}\frac{\sigma_{d,n}^2}{\sigma_{z,n}^2}$, respectively. $\sigma_{d,n}^2$ and $\sigma_{v,n}^2$ are the variances of $u_n(l)w_0$ and $v_n(l)$. $\sigma_{z,n}^2 = 1000E[(u_n(l)w_0)^2]$. The kernel parameter β is 15, $\alpha = 0.5$, $\varepsilon = \varepsilon_1 = \varepsilon_2 = 0.01$.

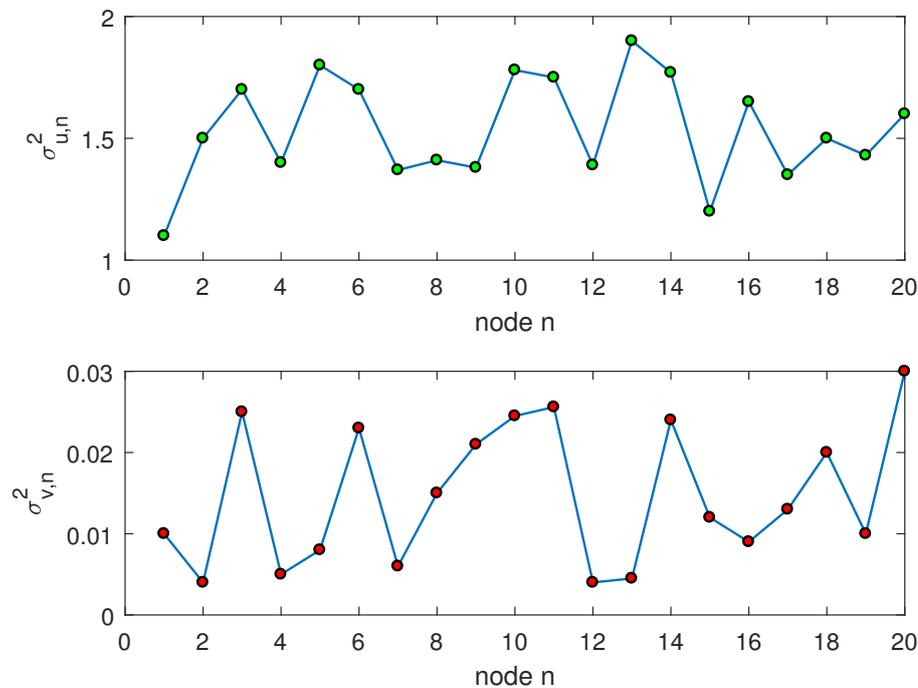


Figure 4. The variance of input signal and Gaussian noise. (up) $\sigma_{u,n}^2$; (down) $\sigma_{v,n}^2$.

Figure 5 gives the performance of the DSAF, MCC-DSAF, and MCC-IPDSAF for colored input signals, where $P_r = 0.01$. Their step-size parameters are 0.1, 0.0158 and 0.0075 to get nearly the same initial convergence. The DSAF is severely degraded under non-Gaussian noise interference. The proposed MCC-DSAF can better suppress non-Gaussian interferences. The steady-state error of MCC-IPDSAF is smaller than MCC-DSAF, which is attributed to the adaptive gain matrix that reassigns the gains to each coefficient.

Figure 6 illustrates the performance of diffusion sign error-LMS (DSE-LMS) [44], DMCC [21], diffusion affine projection sign algorithm (DAPSA) [45], MCC-DSAF, and MCC-IPDSAF for colored input signals for $P_r = 0.01$. Their step-size parameters are 0.0017, 0.065, 0.11, 0.029, and 0.015. The Gaussian kernel $\sigma_{DMCC} = 2$ and the DAPSA projection order is 4. DAPSA converges faster than the DSE-LMS and DMCC, but it converges slower than the proposed MCC-DSAF. MCC-DSAF's convergence is significantly faster than DSE-LMS, DMCC, and DAPSA. It can be verified that the subband algorithms can speed-up the convergence. When the proposed MCC-IPDSAF maintains the same convergence speed with MCC-DSAF, its steady-state error is smaller than that of the MCC-DSAF.

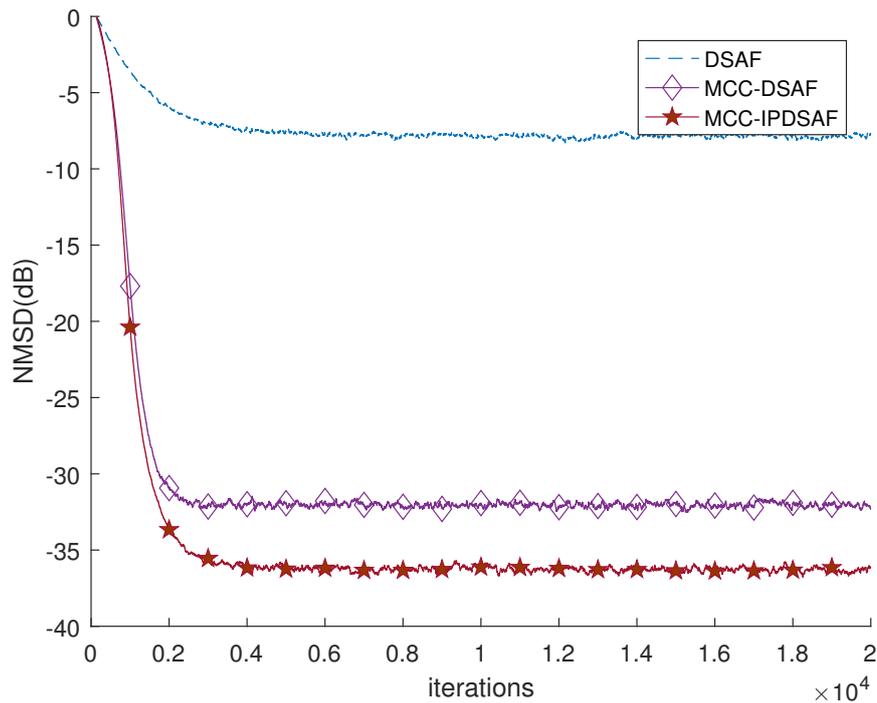


Figure 5. The performance of DSAF, maximum correntropy criterion (MCC)-DSAF, and improved proportionate MCC-DSAF (MCC-IPDSAF) for the colored input with $P_r = 0.01$.

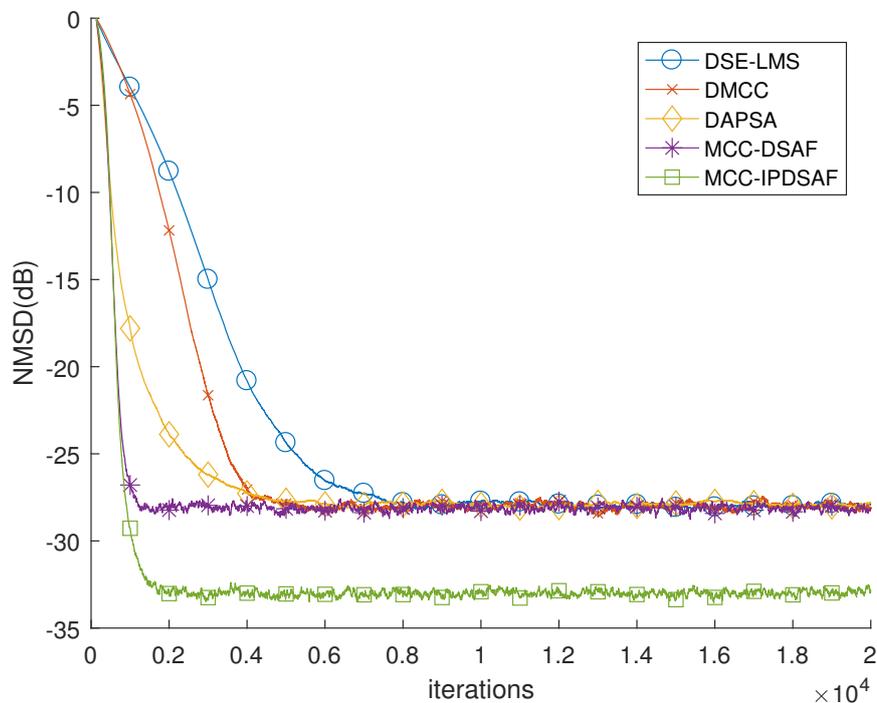


Figure 6. The performance of DSE-LMS, diffusion maximum correntropy criterion (DMCC), DAPSA, MCC-DSAF and MCC-IPDSAF for the colored input with $P_r = 0.01$.

Figure 7 demonstrates the performance of DSAF, DSSAF, IWF-DSSAF, MCC-DSAF, IPDSSAF, IWF-IPDSSAF, and MCC-IPDSAF for a white input signal, where $P_r = 0.01$. Their step-size parameters are chosen to be 0.5, 0.14, 0.07, 0.02, 0.08, 0.055, and 0.01 for achieving the same initial convergence.

DSAF shows the worst behavior. The DSSAF and IWF-DSSAF algorithms have almost the same performance as each other, and the IPDSSAF and IWF-IPDSSAF algorithms are similar, this is because IWF-DSSAF and IWF-IPDSSAF are subband variants of DSSAF and IPDSSAF, respectively, and subband splitting of white signals has almost no effect when the impulsiveness of non-Gaussian noise is not particularly large. The steady-state error of the proposed MCC-DSAF is smaller than the DSSAF and IPDSSAF because it is realized based on MCC, which can resist non-Gaussian noise. As a result, the behavior of MCC-IPDSAF is better than the other algorithms.

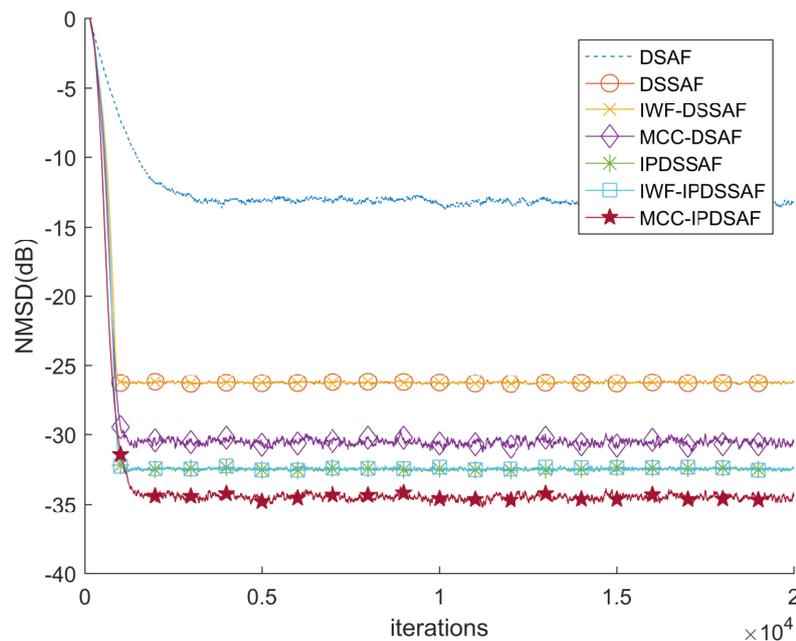


Figure 7. The performance of the DSAF, diffusion sign subband AF algorithm (DSSAF), individual weighting factored (IWF)-DSSAF, MCC-DSAF, IPDSSAF, IWF-IPDSSAF and MCC-IPDSAF for white input with $P_r = 0.01$.

Figure 8 is the performance of DSAF, DSSAF, IWF-DSSAF, MCC-DSAF, IPDSSAF, IWF-IPDSSAF, and MCC-IPDSAF for a white input signal with $P_r = 0.1$. Their step-size parameters are 0.25, 0.7, 0.34, 0.065, 0.4, 0.25, and 0.039, respectively. By increasing the impulsiveness of non-Gaussian noise, DSAF has completely failed, and the algorithms have similar performance to those in Figure 7.

Figure 9 discusses the behavior of DSAF, DSSAF, IWF-DSSAF, MCC-DSAF, IPDSSAF, IWF-IPDSSAF, and MCC-IPDSAF for a colored input signal, where $P_r = 0.01$. The step-size parameters of these algorithms are 0.1, 0.42, 0.055, 0.0158, 0.3, 0.06, and 0.0075, respectively. The DSAF has been severely degraded, while the proposed MCC-DSAF has a smaller steady-state-error than those of DSSAF, IWF-DSSAF, and IPDSSAF. Due to the advantages of adaptive gain matrix and MCC schemes, the proposed MCC-IPDSAF outperforms all other algorithms.

Figure 10 shows the performance of DSAF, DSSAF, IWF-DSSAF, MCC-DSAF, IPDSSAF, IWF-IPDSSAF, and MCC-IPDSAF for a colored input signal, where $P_r = 0.1$. The step size parameters of these algorithms are 0.1, 0.25, 0.06, 0.021, 0.25, 0.0765, and 0.01, respectively. Compared with Figure 9, when the impulsiveness of non-Gaussian noise is increased, the steady-state error and convergence for all algorithms become worse. However, the behavior of the proposed MCC-IPDSAF algorithm is the best.

Figure 11 shows the tracking behavior of DSAF, DSSAF, IWF-DSSAF, MCC-DSAF, IPDSSAF, MCC-IPDSAF, IWF-IPDSSAF for a colored input signal, where $P_r = 0.01$. The unknown system changes when the iterations reach to 20,000, the proposed MCC-DSAF and MCC-IPDSAF still have good tracking performance.

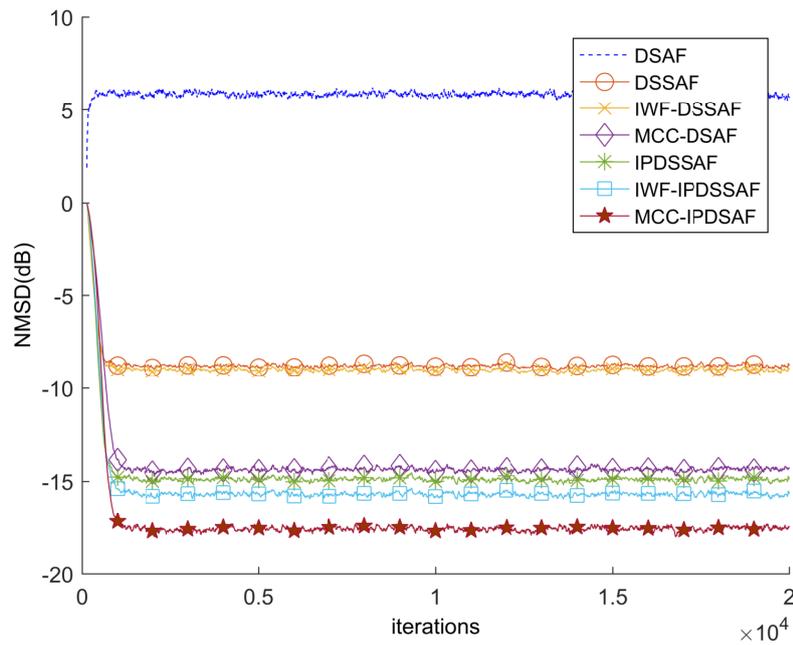


Figure 8. The performance of DSAF, IWF-DSSAF, DSSAF, MCC-DSSAF, IPDSSAF, IWF-IPDSSAF, and MCC-IPDSSAF for white input with $P_r = 0.1$.

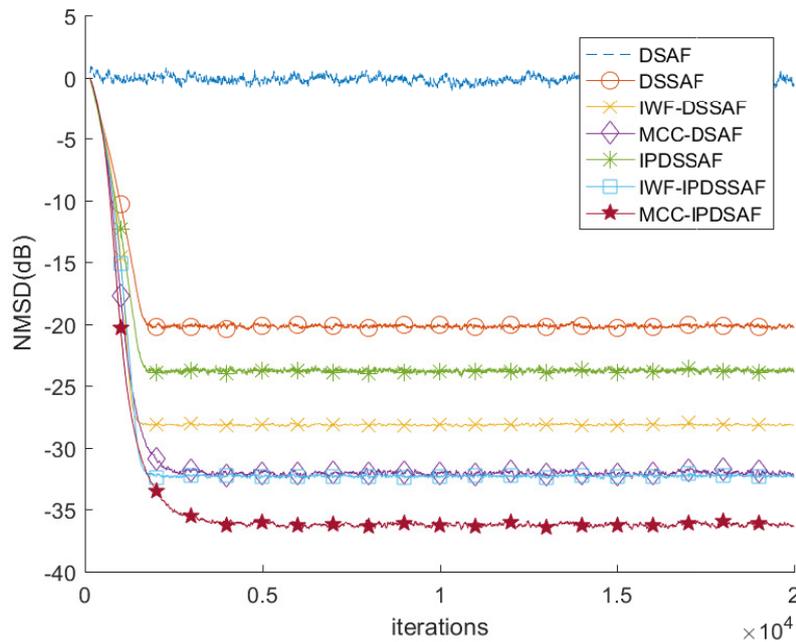


Figure 9. The performance of IWF-DSSAF, DSSAF, DSAF, MCC-DSSAF, IPDSSAF, IWF-IPDSSAF, and MCC-IPDSSAF for the colored input with $P_r = 0.01$.

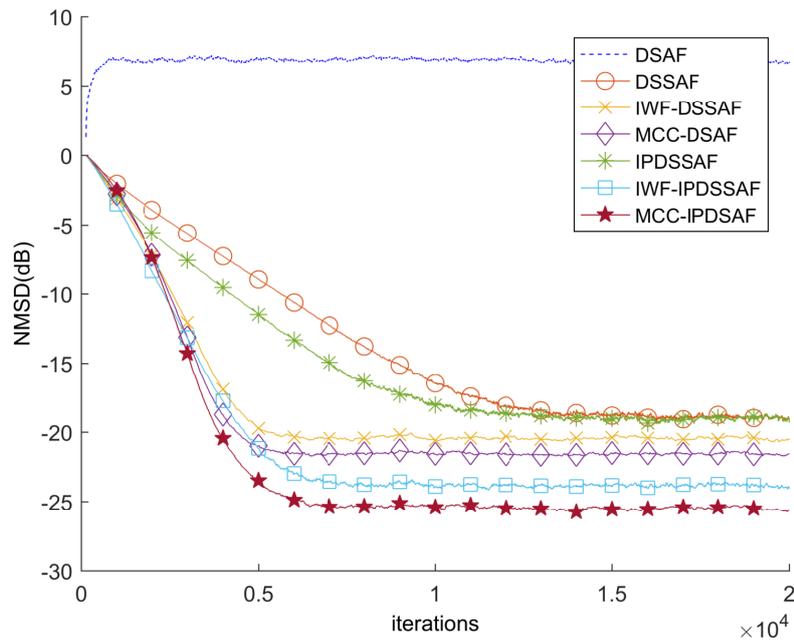


Figure 10. The performance of DSAF, IWF-DSSAF, DSSAF, MCC-DSSAF, IPDSSAF, IWF-IPDSSAF, and MCC-IPDSSAF for the colored input with $P_r = 0.1$.

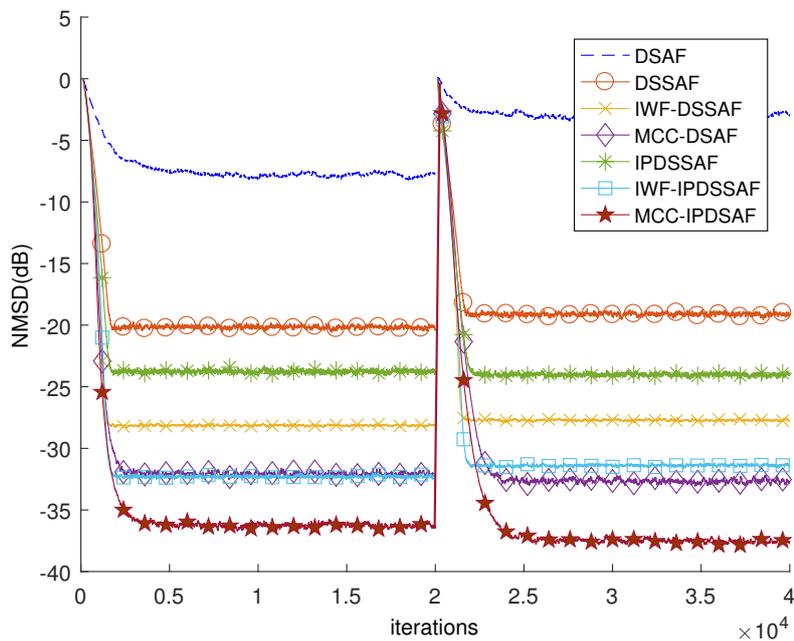


Figure 11. The tracking performance of DSAF, IWF-DSSAF, DSSAF, MCC-DSSAF, IPDSSAF, IWF-IPDSSAF, and MCC-IPDSSAF for the colored input with $P_r = 0.01$.

Figure 12 gives a highly correlated real speech signal and the sparse channel. The real speech signal is the input, and the sampling frequency is 8 KHz, and the sample length is 4.8×10^4 .

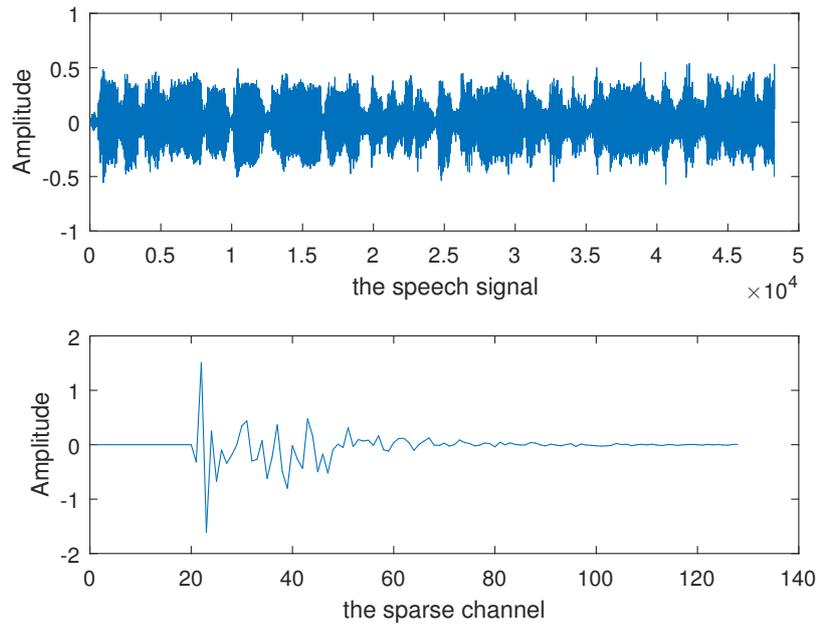


Figure 12. The speech input signal and the sparse channel.

From Figure 13, DSAF fails to converge, while the steady state error of the proposed MCC-DSAF is better than those of the DSSAF and IWF-DSSAF for the non-Gaussian interference and speech input. The behavior of MCC-IPDSAF is superior to the other algorithms. Although the non-stationarity of speech input affects the behavior of the mentioned algorithms, the experiment result verifies the feasibility and effectiveness of the MCC-DSAF and MCC-IPDSAF algorithms.

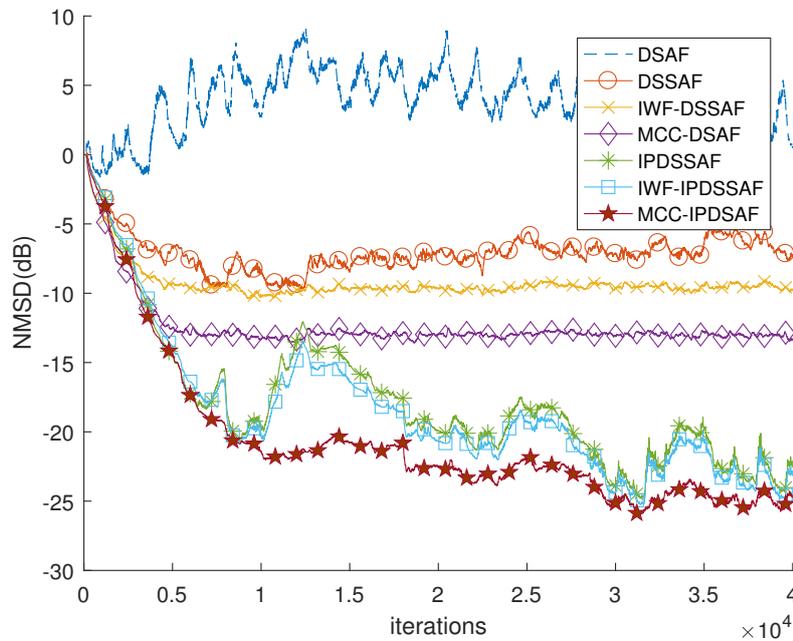


Figure 13. The behavior of DSAF, IWF-DSSAF, DSSAF, MCC-DSAF, IWF-DSSAF, IPDSSAF, and MCC-IPDSAF for speech input with $P_r = 0.01$.

5. Conclusions

In this paper, the maximum correntropy criterion and an individual weighting factor have been taken to construct a new cost function within the distributed subband adaptive filtering framework, which is named MCC-DSAF. The developed MCC-DSAF has been well derived and analyzed. The proposed MCC-DSAF algorithm can not only effectively suppress non-Gaussian noise interference, but also outperforms DSSAF and IWF-DSSAF with respect to convergence and MSD. Moreover, the proportionate adaption scheme is also introduced into MCC-DSAF to get MCC-IPDSAF, which further enhances the behavior of MCC-DSAF for identifying sparse systems. The convergence analysis and the steady-state behavior of MCC-DSAF are presented. The estimation behaviors of the algorithms are verified and the simulation results demonstrate that the proposed MCC-DSAF and MCC-IPDSAF are superior to the mentioned, popular DSAF algorithms. The algorithms in this paper will provide a better effect in the fields of radar, medical, wireless sensor networks, smart dust networks, distributed channel estimations, and hydroacoustics, etc.

Author Contributions: Y.G. wrote the draft of this paper and gave the idea and basic simulation of the devised algorithm. J.L. proposed the analysis of the algorithms and constructed the figures with other algorithms. Y.L. gave the idea and the investigation of the devised algorithm. All the authors discuss the results and the analysis of the devised algorithms together, and they have read and agreed to submit the final version of the manuscript.

Funding: This paper is supported by the National Key Research and Development Program of China (2016YFE0111100), Key Research and Development Program of Heilongjiang (GX17A016), the Science and Technology innovative Talents Foundation of Harbin (2016RAXXJ044), the Natural Science Foundation of Beijing (4182077), China Postdoctoral Science Foundation (2017M620918, 2019T120134), the Fundamental Research Funds for the Central Universities (HEUCFG201829, 2072019CFG0801), and Opening Fund of Acoustics Science and Technology Laboratory (SSKF2016001).

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

References

1. Sayed, A.H. Adaptive, networks. *Procs. IEEE* **2014**, *102*, 460–497. [[CrossRef](#)]
2. Rossi, L.A.; Krishnamachai, B.; Kuo, C.C.J. Distributed parameter estimation for monitoring diffusion phenomena using physical models. In Proceedings of the 2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks, Santa Clara, CA, USA, 4–7 October 2004.
3. Li, J.; AlRegib, G. Rate-constrained distributed estimation in wireless sensor network. *IEEE Trans. Signal Process.* **2007**, *55*, 1624–1643. [[CrossRef](#)]
4. Culler, D.; Estrin, D.; Srivastava, M. Overview of sensor networks. *Computer* **2004**, *37*, 41–49. [[CrossRef](#)]
5. Tu, S.Y.; Sayed, A.H. Mobile adaptive networks. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 649–664.
6. Sayed, A.H. Adaptive, learning, and optimization over networks. *Found. Trends Mach. Learn.* **2014**, *7*, 311–801. [[CrossRef](#)]
7. Chen, F.; Shi, T.; Duan, S.; Wang, L.; Wu, J. Diffusion Least Logarithmic Absolute Difference Algorithm for Distributed estimation. *Signal Process.* **2018**, *142*, 423–430. [[CrossRef](#)]
8. Li, L.; He, J.; Zhang, Y. Affine projection algorithms over distributed wireless networks. *J. Commun. Univ. China (Sci. Technol.)* **2012**, *19*, 34–43.
9. 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](#)]
10. Jin-gen, N.I.; Lan-shen, M.A. Distributed subband adaptive filtering algorithms. *Chin. J. Electron.* **2015**, *43*, 2222–2231.
11. Seo, J.H.; Jung, S.M.; Park, P.G. A diffusion subband adaptive filtering algorithm for distributed estimation using variable step size and new combination method based on the MSD. *Digit. Signal Process.* **2016**, *48*, 361–369. [[CrossRef](#)]
12. Abadi, M.S.E.; Ahmadi, M.J. Diffusion Improved Multiband-Structured Subband Adaptive Filter Algorithm with Dynamic Selection of Nodes Over Distributed Networks. *IEEE Trans. Circuits Syst.* **2019**, *66*, 507–511. [[CrossRef](#)]
13. Abadi, M.S.E.; Shafiee, M.S. A family of diffusion normalized subband adaptive filter algorithms over distributed networks. *Int. J. Commun. Syst.* **2017**, *30*, 1–15.

14. Abadi, M.S.E.; Shafiee, M.S. Diffusion Normalized Subband Adaptive Algorithm for Distributed Estimation Employing Signed Regressor of Input Signal. *Digit. Signal Process.* **2017**, *70*, 73–83. [[CrossRef](#)]
15. Wen, P.; Zhang, J. Widely-Linear Complex-Valued Diffusion Subband Adaptive Filter Algorithm. *IEEE Trans. Signal Inf. Process. Netw.* **2018**, *47*, 90–95. [[CrossRef](#)]
16. Mathews, V.J.; Cho, S.H. Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **1987**, *35*, 450–454. [[CrossRef](#)]
17. Ni, J. Diffusion sign subband adaptive filtering algorithm for distributed estimation. *IEEE Signal Process. Lett.* **2015**, *11*, 2029–2033. [[CrossRef](#)]
18. Shi, J.; Ni, J. Diffusion sign subband adaptive filtering algorithm with enlarged cooperation and its variant. *Circuits Syst. Signal Process.* **2016**, *36*, 1714–1724. [[CrossRef](#)]
19. Wang, W.Y.; Zhao, H.Q.; Liu, Q.Q. Diffusion Sign Subband Adaptive Filtering Algorithm with Individual Weighting Factors for Distributed Estimation. *Circuits Syst. Signal Process.* **2017**, *36*, 2605–2621. [[CrossRef](#)]
20. Chen, B.; Principe, J.C. Maximum correntropy estimation is a smoothed MAP estimation. *IEEE Signal Process. Lett.* **2012**, *19*, 491–494. [[CrossRef](#)]
21. Ma, W.; Chen, B.; Duan, J.; Zhao, H. Diffusion maximum correntropy criterion algorithms for robust distributed estimation. *Digit. Signal Process.* **2016**, *58*, 10–19. [[CrossRef](#)]
22. Li, C.; Shen, P.; Liu, Y. Diffusion information theoretic learning for distributed estimation over network. *IEEE Trans. Signal Process.* **2013**, *61*, 4011–4024. [[CrossRef](#)]
23. Singh, A.; Principe, J.C. Using correntropy as a cost function in linear adaptive filters. In Proceedings of the International Joint Conference, Neural Networks (IJCNN), Atlanta, GA, USA, 14–19 June 2009; pp. 2950–2955.
24. Li, Y.; Jiang, Z.; Shi, W.; Han, X.; Chen, B. Blocked maximum correntropy criterion algorithm for cluster-sparse system identifications. *IEEE Trans. Circuits Syst. II Exp. Briefs* **2019**. [[CrossRef](#)]
25. Shi, W.; Li, Y.; Wang, Y. Noise-free maximum correntropy criterion algorithm in non-Gaussian environment. *IEEE Trans. Circuits Syst. II Exp. Briefs* **2019**. [[CrossRef](#)]
26. Haykin, S.; Widrom, B. *Least-Mean-Square Adaptive Filters*. Wiley: Hoboken, NJ, USA, 2005.
27. Li, Y.; Wang, Y.; Jiang, T. Sparse least mean mixed-norm adaptive filtering algorithms for sparse channel estimation applications. *Int. J. Commun. Syst.* **2017**, *30*, e3181. [[CrossRef](#)]
28. Li, Y.; Wang, Y.; Jiang, T. Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation. *Signal Process.* **2016**, *128*, 243–251. [[CrossRef](#)]
29. Li, Y.; Jiang, Z.; Osman, O.; Han, X.; Yin, J. Mixed norm constrained sparse APA algorithm for satellite and network echo channel estimation. *IEEE Access* **2018**, *6*, 65901–65908. [[CrossRef](#)]
30. Shi, W.; Li, Y.; Zhao, L.; Liu, X. Controllable sparse antenna array for adaptive beamforming. *IEEE Access* **2019**, *7*, 6412–6423. [[CrossRef](#)]
31. Zhang, X.; Jiang, T.; Li, Y.; Zakharov, Y. A novel block sparse reconstruction method for DOA estimation with unknown mutual coupling. *IEEE Commun. Lett.* **2019**, *23*, 1845–1848. [[CrossRef](#)]
32. Jin, Z.; Li, Y.; Wang, Y. An enhanced set-membership PNLMS algorithm with a correntropy induced metric constraint for acoustic channel estimation. *Entropy* **2017**, *19*, 281. [[CrossRef](#)]
33. Guo, Y.; Gao, Y. Comparative Study of Several Sparse Adaptive Filtering Algorithms. *Microcomput. Appl.* **2014**, *33*, 1–3.
34. Shi, L.; Zhao, H.Q. Two Diffusion Proportionate Sign Subband Adaptive Filtering Algorithms. *Circuits Syst. Signal Process.* **2017**, *36*, 4244–4259. [[CrossRef](#)]
35. Chen, B.; Xing, L.; Liang, J.; Zheng, N.; Principe, J.C. Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion. *IEEE Signal Process. Lett.* **2014**, *21*, 880–884.
36. Wu, Q.; Li, Y.; Zakharov, Y.; Xue, W.; Shi, W. A kernel affine projection-like algorithm in reproducing kernel hilbert space. *IEEE Trans. Circuits Syst. II Exp. Briefs* **2019**. [[CrossRef](#)]
37. Ma, W.T.; Duan, J.D.; Man, W.S.; Liang, J.L.; Chen, B.D. General Mixed Norm Based Diffusion Adaptive Filtering Algorithm for Distributed Estimation over Network. *IEEE Access* **2017**, *5*, 1090–1102. [[CrossRef](#)]
38. Saeed, M.O.B.; Zerguine, A.; Zummo, S.A. A variable step-size strategy for distributed estimation over adaptive networks. *EURASIP J. Adv. Signal Process.* **2013**, *2013*, 135. [[CrossRef](#)]
39. Xu, X.G.; Qu, H.; Zhao, J.H.; Yan, F.Y.; Wang, W.H. Diffusion Maximum Correntropy Criterion Based Robust Spectrum Sensing in Non-Gaussian Noise Environments. *Entropy* **2018**, *20*, 246. [[CrossRef](#)]

40. Tracy, D.S.; Singh, R.P. A new matrix product and its applications in partitioned matrix differentiation. *Stat. Neerl.* **1972**, *26*, 143–157. [[CrossRef](#)]
41. Cattivelli, F.; Sayed, A.H. Diffusion LMS Strategies for Distributed Estimation. *IEEE Trans. Signal Process.* **2010**, *58*, 1035–1048. [[CrossRef](#)]
42. G, Y.; B, Y.M. Robust Affine Projection Sign Adaptive Filtering Algorithm. *J. Instrum. Instrum.* **2017**, *38*, 23–32.
43. Li, Y.; Wang, Y.; Jiang, T. Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo cancelation. *AEU Int. J. Electron. Commun.* **2016**, *70*, 895–902. [[CrossRef](#)]
44. Ni, J.; Chen, J.; Chen, X. Diffusion sign-error LMS algorithm: Formulation and stochastic behavior analysis. *Signal Process.* **2016**, *128*, 142–149. [[CrossRef](#)]
45. Ni, J.; Ma, L. Distributed affine projection algorithms against impulsive interferences. *Tien Tzu Hsueh Pao/Acta Electron. Sin.* **2016**, *44*, 1555–1560



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).