IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 3, 2019, accepted June 17, 2019, date of publication June 20, 2019, date of current version October 30, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2924028

Active Coefficient Detection Maximum Correntropy Criterion Algorithm for Sparse Channel Estimation Under Non-Gaussian Environments

ZEYANG SUN[1], YINGSONG LI[1,2], (Senior Member, IEEE), ZHENGXIONG JIANG[1], AND WANLU SHI[1]

[1]College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China
[2]Key Laboratory of Microwave Remote Sensing, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

Corresponding author: Yingsong Li (liyingsong@ieee.org)

This work was supported in part by the Natural Science Foundation of Beijing under Grant 4182077, in part by the Fundamental Research Funds for the Central Universities under Grant 3072019CFG0801, and in part by the Opening Fund of Acoustics Science and Technology Laboratory under Grant SSKF2016001.

ABSTRACT In this paper, a kind of active coefficient detection (ACD)-based maximum correntropy criterion (MCC) algorithm is proposed to estimate a sparse multi-path channel under the non-Gaussian environments. The proposed ACD-based MCC algorithms are realized by developing an active coefficient detection mechanism, which can distinguish the active taps within the sparse channels and find out the position and the number of active taps. Therefore, only the active taps coefficient is updated while the trivial channel coefficients are set to be zeros. Various computer simulation experiments are carried out to investigate the performance of the proposed ACD-based MCC algorithms under different impulsive noises. The achieved simulation results prove that the proposed ACD-based MCC algorithms are effective and outperform the previous adaptive filtering algorithms for the sparse channel estimation with regard to both the convergence and the estimation error.

INDEX TERMS Sparse channel estimation, maximum correntropy criterion, active coefficient detection, tap selection, impulsive noise environments.

I. INTRODUCTION

With the growth of wireless communication technologies, broadband information transmission has become a hot topic to meet the demand of users [1]. Thus, broadband communication has attracted more attention for the researchers domestic and overseas. In the wireless communication transmission, the frequency-selective channel fading is inevitable, and therefore, channel estimation has become an amazing research topic to get high quality. Then, various effective channel estimation methods are reported in [2]–[8], [10]–[12] to enhance the accuracy of channel state information estimation. In fact, for tracking the channel in the practical systems, the adaptive learning techniques are proposed in the adaptive filtering framework and deep learning directions [13]–[15]. Among these methods, [8]–[12] provided good performance in adaptive channel estimations (ACEs) and low complexity. Moreover, least mean square (LMS) algorithm is one of the most classical adaptive filter (AF) algorithms, which has been used for ACE in single-input single-output systems like orthogonal frequency-division multiplexing (OFDM) [16], [17] and self-cancellation in full duplex communication systems [18], [19]. After that, a series of effective ACE methods have been proposed in the past decades [2]–[6], [8]–[12].

On the other hand, the broadband communication channel is always regarded as the sparse channel [8]–[12], in which most of the channel response coefficients are zero or near zero that indicate the channel taps are inactive. To utilize the inherent sparse structure properties in such channels, several ACE algorithms have been proposed and derived within the AF framework [8]–[12] to estimate these channels. For example, the proportionate NLMS (PNLMS) algorithm takes advantage of the sparsity in the unknown channel by

The associate editor coordinating the review of this manuscript and approving it for publication was Guan Gui.

allocating different gains to the channel coefficients. According to the proportionate theory in [20], after that, numerous PNLMS-based AF methods [21]–[24] have been reported with an improved estimation performance.

Furthermore, in recent years, another kind of sparse ACE methods [8], [10]–[12] have been put forward on the basis of the zero-attracting (ZA) scheme which is enlightened by the recent invention in compressive sensing (CS) [25]. Based on this idea, the ZA-LMS and the reweighted ZA-LMS (RZA-LMS) algorithms [26] have been put forward via integrating a sparsity-aware penalty on channel coefficient vector. Then, many enhanced sparse ACE algorithms have been further studied using various norm penalties [27], [28]. Although these sparse AF algorithms can improve channel estimation behavior, most of them are implemented based on the Gaussian model. However, from channel measurement results, we found that the non-Gaussian model is more suitable than the Gaussian model to describe the real communication environment since the existing impulse noises and manual noises in the systems [30]. Furthermore, the impulsive noise is common in the real world, and then, the estimation performance might degrade when the signal is transmitted from the transmitter to receiver because of the noise contamination [29]–[31].

To resist the impulsive noises, the maximum correntropy criterion (MCC) has been brought out to combat against non-Gaussian noise in channel estimation applications, where the correntropy is a smooth nonlinear correlation measure [32] which reveals the correlation between two stochastic variables [33], [34]. The MCC-based algorithm has the outstanding capability in the presence of impulsive noise scenarios [35], [36] in comparison with the LMS-based algorithms. Consequently, correntropy theory has made a significant contribution to the application in robust regression in AFs [37], spectral characterization [38], and so on. However, the MCC algorithm cannot utilize the sparse structure characteristics of the sparse systems [39], [40]. Inspired by the ZA-LMS and the RZA-LMS algorithms, the ZA-MCC and RZA-MCC ACE algorithms have been presented in [30] to improve the steady-state behavior and convergence of the MCC algorithm for estimating multi-path channels.

In this paper, the active coefficient detection based maximum correntropy criterion (ACD-based MCC) algorithms which consist of ACD-MCC algorithm and ACD-normalized MCC (ACD-NMCC) algorithm are proposed and discussed to develop a sparse ACE for estimating sparse channels under the impulsive noises. The proposed ACD-based MCC algorithms are realized and implemented by utilizing an ACD criterion which improves the ability to predict the non-zero coefficients in the unknown channels. In particular, the active coefficients detection (ACD) criterion can effectively find out and update the active taps, and hence, the ACD-based algorithms can get an accelerated convergence rate and mitigate the impulsive noise interferences. Numerical experiments indicate that the proposed ACD-based MCC algorithms can obtain faster convergence and better steady-state performance for different input signals and various impulsive environments.

The rest of this paper consists of the following sections. In section II, the correntropy theory and MCC are briefly reviewed. In section III, the ACD scheme and ACD-based MCC algorithms are carefully derived and analyzed. The simulation results are presented and discussed in section IV, and a conclusion is drawn in section V.

II. REVIEW OF THE MCC ALGORITHM

A. CORRENTROPY

Correntropy is to effectively describe the similarity between two arbitrary random variables X and Y with same dimension, which is defined as [41], [42]

$$V(X,Y) = E[\kappa(X,Y)] = \int \kappa(x,y)dF_{XY}(x,y), \tag{1}$$

where E is an expectation operator, $\kappa(.,.)$ is a shift-invariant Mercer kernel whose kernel width is determined by parameter σ, and $F_{XY}(x,y)$ denotes the joint distribution function of the random variables X and Y. In fact, the joint probability density function is unknown, and only a finite number of samples $\{(x_i,y_i)\}_{i=1}^{N}$ are available. Thus, we can obtain the sample estimator [40]

$$\hat{V}_N(X,Y) = \frac{1}{N}\sum_{i=1}^{N}\kappa(x_i,y_i). \tag{2}$$

Here, the most widely used kernel, normalized Gaussian kernel G_σ with variance of σ [42], [43], is used and given by

$$\kappa(x,y) = G_\sigma(x-y) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{(x-y)^2}{2\sigma^2}). \tag{3}$$

Thus, the correntropy is obtained [45]

$$\begin{aligned}\hat{V}(X,Y) &= \frac{1}{\sqrt{2\pi}\sigma}E[\exp\left(\frac{-(X-Y)^2}{2\sigma^2}\right)] \\ &= \frac{1}{\sqrt{2\pi}\sigma}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n n!}E\left[\frac{(X-Y)^{2n}}{\sigma^{2n}}\right].\end{aligned} \tag{4}$$

B. MCC ALGORITHM

To get the derivation of the MCC algorithm, the transcendental knowledge of the AF illustrated in [44], [45] is considered in this paper. $\mathbf{X}(n)$ represents the input signal vector at the sample instant n, $\mathbf{W}$ is the channel impulse response (CIR) of the unknown channel, $v(n)$ is the measurement noises existing in the unknown channel. The gotten signal $r(n)$ can be described as

$$r(n) = \mathbf{X}^T(n)\mathbf{W} + v(n), \tag{5}$$

where $\mathbf{X}(n) = [x(n), x(n-1), \ldots, x(n-M+1)]^T$ and $\mathbf{W} = [w_0, w_1, \ldots, w_{M-1}]^T$. M is the dimension of the unknown system. On the basis of the ACE, the output signal $y(n)$ of the estimator is

$$y(n) = \mathbf{X}^T(n)\hat{\mathbf{W}}(n-1), \tag{6}$$

where $\hat{\mathbf{W}}(n)$ is an estimated channel coefficient vector given by $\hat{\mathbf{W}}(n) = [\hat{\omega}_0(n), \hat{\omega}_1(n), \ldots, \hat{\omega}_{M-1}(n)]^T$. Thus, the estimation error is

$$e(n) = r(n) - y(n) = r(n) - \mathbf{X}^T(n)\hat{\mathbf{W}}(n). \tag{7}$$

The correntropy is used to construct a cost function for the MCC algorithm [46], [47] to solve the following equation

$$\begin{aligned} J_n &= \frac{1}{\sqrt{2\pi}\sigma}\cdot\frac{1}{N}\sum_{i=n-N+1}^{n}\exp\left(\frac{-(r(i)-y(i))^2}{2\sigma^2}\right) \\ &= \frac{1}{\sqrt{2\pi}\sigma}\cdot\frac{1}{N}\sum_{i=n-N+1}^{n}\exp\left(\frac{-e^2(i)}{2\sigma^2}\right) \end{aligned} \tag{8}$$

Then, iterative gradient ascent approach is considered for getting the updating equation of filter coefficients [48]. Thus, we have

$$\hat{\mathbf{W}}(n+1) = \hat{\mathbf{W}}(n) + \varepsilon\nabla J(n), \tag{9}$$

where ε is the step size. Considering (8) and (9), we get

$$\begin{aligned} \hat{\mathbf{W}}(n+1) = \hat{\mathbf{W}}(n) &+ \frac{\varepsilon}{\sqrt{2\pi}\sigma^3}\frac{1}{N} \\ &\times \sum_{i=n-N+1}^{n}\left[\exp\left(\frac{-e^2(i)}{2\sigma^2}\right)e(i)x(i)\right]. \end{aligned} \tag{10}$$

Based on (10), the vector form of the MCC's iterative equation is given by [49],

$$\hat{\mathbf{W}}(n+1) = \hat{\mathbf{W}}(n) + \frac{\varepsilon}{\sqrt{2\pi}\sigma^3}\exp\left(\frac{-e^2(n)}{2\sigma^2}\right)e(n)\mathbf{X}(n). \tag{11}$$

Simplifying (11), the adaptive process of the tap weight vector can be expressed as [50]

$$\hat{\mathbf{W}}(n+1) = \hat{\mathbf{W}}(n) + \mu\exp\left(\frac{-e^2(n)}{2\sigma^2}\right)e(n)\mathbf{X}(n), \tag{12}$$

where $\mu = \frac{\varepsilon}{\sqrt{2\pi}\sigma^3}$ is the renewed step-size for basic MCC algorithm. Similarity, inspired by the normalized LMS (NLMS) algorithm, the NMCC algorithm has been reported to enhance the stability of the MCC algorithm, whose updating equation is [51]

$$\hat{\mathbf{W}}(n+1) = \hat{\mathbf{W}}(n) + \mu\frac{\exp\left(\frac{-e^2(n)}{2\sigma^2}\right)}{||\mathbf{X}(n)||^2 + \delta}e(n)\mathbf{X}(n), \tag{13}$$

where δ is a regularization factor which is to guarantee that the denominator of (13) is not being zero. From the NMCC, it is noted that there is an extra exponential term in the MCC-based algorithms for the sake of comparison with the LMS algorithm to resist the impulse noise.

III. THE PROPOSED ACD-BASED MCC ALGORITHMS

The ACD scheme is devised to seek the positions and number of the active taps, which can accelerate the speed of MCC algorithm. To address the ACD-based MCC algorithms, a filter coefficient detection criterion is illustrated by considering the least-square (LS) method if the signal and the unknown channel satisfies the following restrictions.

Restriction 1: Both the input signal and the interference signal are zero mean, bounded and generalized stationary processes, and they are uncorrelated to each other over time.

Restriction 2: The autocorrelation matrix of the input signal is positive definite and can denote as

$$\mathbf{R}_{xx} = \sigma_x^2\mathbf{I},$$

where σ_x^2 is the power of input signal and $\mathbf{I}$ is a $M \times M$ identity matrix.

Restriction 3: The unknown system has only k active taps within its length. The objective is to get the estimation of $\mathbf{W}$ by solving the following cost function

$$\begin{aligned} \zeta_N(\hat{\mathbf{W}}(n)) &\triangleq \sum_{n=1}^{N}(r(n) - y(n))^2 \\ &= \sum_{n=1}^{N}\left(r(n) - \mathbf{X}^T(n)\hat{\mathbf{W}}(n)\right)^2. \end{aligned} \tag{14}$$

To further explain the ACD scheme based on the filter coefficient detection criterion, the optimal wiener solution is used to seek the solution, which is

$$\hat{\mathbf{W}}^{LS}(N) = \mathbf{R}^{-1}(N)\mathbf{P}(N), \tag{15}$$

where $\mathbf{P}(N)$ is the cross-correlation vector between the filter input and desired signal, and $\mathbf{R}(N)$ is the input signal autocorrelation matrix [52]. Thereby, we have

$$\mathbf{P}(N) = \frac{1}{N}\sum_{n=1}^{N}r(n)\mathbf{X}(n), \tag{16}$$

$$\mathbf{R}(N) = \frac{1}{N}\sum_{n=1}^{N}\mathbf{X}(n)\mathbf{X}^T(n). \tag{17}$$

Assuming that there are k nonzero CIR elements in $\mathbf{W}$, their positions c_j constitute a specific array $\mathbf{c}^k$, where j represents the j-th active tap.

$$\mathbf{c}^k = \{c_1, c_2, \ldots, c_k\},$$

where $0 \le c_j < M$, $j = 1, 2, \ldots, k$. At iteration N, the estimated active tap coefficients is given by

$$\hat{\mathbf{W}}(N, \mathbf{c}^k) \triangleq [\hat{\omega}_{c_1}(N), \hat{\omega}_{c_2}(N), \ldots, \hat{\omega}_{c_k}(N)]^T. \tag{18}$$

From (18), to further obtain the channel coefficient activity, a $M \times k$ matrix $\mathbf{H}(\mathbf{c}^k)$ is presented, which is composed of $\mathbf{H}(c_j, j) = 1, j = 1, 2, \ldots, k$. Then, the LS's cost function in (14) is rewritten as

$$\begin{aligned} \zeta_N(\hat{\mathbf{W}}(N, \mathbf{c}^k)) = N&\left[\mathbf{H}(\mathbf{c}^k)\hat{\mathbf{W}}(N, \mathbf{c}^k) - \hat{\mathbf{W}}^{\mathrm{LS}}(N)\right]^T \\ &\times\mathbf{R}(N)\left[\mathbf{H}(\mathbf{c}^k)\hat{\mathbf{W}}(N, \mathbf{c}^k) - \hat{\mathbf{W}}^{\mathrm{LS}}(N)\right] \\ &+\zeta_N\left(\hat{\mathbf{W}}^{\mathrm{LS}}(N)\right), \end{aligned} \tag{19}$$

where $\zeta_N\left(\hat{\mathbf{W}}^{\mathrm{LS}}(N)\right)$ corresponds to the minimum of the LS's cost function in (14) when the optimal winner solution is applied to the LS method. And, $\zeta_N\left(\hat{\mathbf{W}}^{\mathrm{LS}}(N)\right)$ is independent

of the filter coefficient $\hat{\mathbf{W}}$. As a result, equation (19) is modified to be a new cost function with $\hat{\mathbf{W}}$ as the variable, which is given by

$$C_N(\hat{\mathbf{W}}(N,\mathbf{c}^k)) = N\left[\mathbf{H}(\mathbf{c}^k)\hat{\mathbf{W}}(N,\mathbf{c}^k) - \hat{\mathbf{W}}^{\mathrm{LS}}(N)\right]^T \times \mathbf{R}(N)\left[\mathbf{H}(\mathbf{c}^k)\hat{\mathbf{W}}(N,\mathbf{c}^k) - \hat{\mathbf{W}}^{\mathrm{LS}}(N)\right], \tag{20}$$

Next, only k active taps in the optimal LS's solution is considered to get their effects on (20), which is given by [53]

$$\begin{aligned}\hat{\mathbf{W}}^{LS}(N,\mathbf{c}^k) &= [\mathbf{H}^T(\mathbf{c}^k)\mathbf{R}(N)\mathbf{H}(\mathbf{c}^k)]^{-1}\mathbf{H}^T(\mathbf{c}^k)\mathbf{R}(N)\hat{\mathbf{W}}^{LS}(N)\\ &= [\mathbf{H}^T(\mathbf{c}^k)\mathbf{R}(N)\mathbf{H}(\mathbf{c}^k)]^{-1}\mathbf{H}^T(\mathbf{c}^k)\mathbf{P}(N),\end{aligned} \tag{21}$$

and then, (21) is substituted to (20) to get

$$C_N(\hat{\mathbf{W}}^{LS}(N,\mathbf{c}^k)) = N\mathbf{P}^T(N)[\mathbf{R}(N)^{-1} - \mathbf{H}(\mathbf{c}^k) \times [\mathbf{H}^T(\mathbf{c}^k)\mathbf{R}(N)\mathbf{H}(\mathbf{c}^k)]^{-1}\mathbf{H}^T(\mathbf{c}^k)]\mathbf{P}(N) \tag{22}$$

Obviously, the cost function in (22) is only related to the cross-correlation vector $\mathbf{P}(N)$, autocorrelation matrix $\mathbf{R}(N)$ and the position of the active tap $\mathbf{c}^k$. The position of k active taps can be gotten by the set $\mathbf{c}^k$ via minimizing the right hand side (RHS) of (22). However, it is extremely complicated to get the optimal solution in the comparison of $\frac{M!}{k!(M-k)!}$ results.

Using the Lemma 1 in [53], when N is large enough and the input signal is white Gaussian noise, the cost function in (22) can be modified to be

$$\frac{1}{N}||C_N(\hat{\mathbf{W}}^{\mathrm{LS}}(N,\mathbf{c}^k) - \bar{C}_N(\mathbf{c}^k))||_2 \to 0, \quad when \quad N \to \infty \tag{23}$$

In a word, $C_N(\hat{\mathbf{W}}^{\mathrm{LS}}(N,\mathbf{c}^k)$ can be approximated by $\overline{C}_N\left(\mathbf{c}^k\right)$ when N approaches infinity, which are given by

$$\bar{C}_N(\mathbf{c}^k) = N\mathbf{P}^T(N)[\bar{\mathbf{R}}(N)^{-1} - \mathbf{H}(\mathbf{c}^k) \times [\mathbf{H}^T(\mathbf{c}^k)\bar{\mathbf{R}}(N)\mathbf{H}(\mathbf{c}^k)]^{-1}\mathbf{H}(\mathbf{c}^k)^T]\mathbf{P}(N), \tag{24}$$

where $\bar{\mathbf{R}}$ is an $M \times M$ diagonal matrix whose elements are given by $\bar{R}_{pp}(N) = \frac{1}{N}\sum_{n=p}^{N} x(n-p-1)^2$, where $p = 1, 2, 3, \ldots, M$, p represents arbitrary channel taps in $\hat{\mathbf{W}}$. Till now, we can see that each channel coefficient has independent effect on the cost function $\overline{C}_N\left(\mathbf{c}^k\right)$. That is to say the positions of k active taps are determined by the k maximum values, which is obtained by

$$Z(N,p) \triangleq \frac{\left[\sum_{n=p+1}^{N} r(n)x(n-p)\right]^2}{\sum_{n=p+1}^{N} x^2(n-p)}, \tag{25}$$

where p is the arbitrary tap in $\hat{\mathbf{W}}$. Using equation (25), we can get the most active k taps rather than distinguishing the active taps and inactive taps. Consequently, a threshold is used to detect the active taps. To achieve this goal, a threshold is constructed to detect the active taps, and the cost function should satisfy the property that it can be minimized by $\hat{\mathbf{W}}$ that contains only active taps as N approaches infinity. However, the LS's cost function doesn't satisfy this requirement. Then, a new cost function is obtained by adding an extra penalty B(N), which is given by

$$\log\zeta(\hat{\mathbf{W}}) + k\frac{B(N)}{N}, \tag{26}$$

In (26), we have $B(N) \to \infty$, $\frac{B(N)}{N} \to 0$ when $N \to \infty$. Herein, $B(N)$ is obtained by the Akaike's B-Information Criterion [57], and hence, equation (26) changes to be

$$\begin{aligned}\zeta'_N &= \zeta(\hat{\mathbf{W}}) + Qk\log N\\ &= \zeta(\hat{\mathbf{W}})\left[1 + \frac{QNk\log(N)/\zeta(\hat{\mathbf{W}})}{N}\right]\end{aligned} \tag{27}$$

where $Q > 0$ is a constant. By choosing a proper Q, the convergence speed will be greatly improved. To seek the solution, following the procedure from (14) to (25), a single-tap LS cost function is employed and defined as

$$\bar{\zeta}_N(p) \triangleq \sum_{n=1}^{N} r^2(n) - Z(N,p). \tag{28}$$

Utilizing the Lemma 1 in [53], (28) is approximated to be

$$\frac{1}{N}\sum_{j=1}^{k}\bar{\zeta}_N(c_j) \to \frac{1}{N}\zeta_N(\hat{\mathbf{W}}(N,\mathbf{c}^k)) + \frac{k-1}{N}\sum_{n=1}^{N} r^2(n) \quad \text{as} \quad \mathrm{N} \to \infty \tag{29}$$

From (27), we can get

$$\frac{1}{N}\zeta'_N(N,\mathbf{c}^k) \triangleq \frac{1}{N}\zeta_N(\hat{\mathbf{W}}(N,\mathbf{c}^k)) + Qk\frac{\log N}{N}. \tag{30}$$

Combining (29) and (30), we obtain

$$\frac{1}{N}\zeta'_N(\mathbf{c}^k) \triangleq \frac{1}{N}\sum_{j=1}^{k}\bar{\zeta}'_N(c_j) - \frac{k-1}{N}\sum_{n=1}^{N} r^2(n), \tag{31}$$

where

$$\bar{\zeta}'_N(c_j) \triangleq \bar{\zeta}_N(c_j) + Q\log N.$$

Then,

$$\frac{1}{N}\bar{\zeta}'_N(\mathbf{c}^k) = \frac{1}{N}\sum_{n=1}^{N} r^2(n) - \sum_{j=1}^{k}\left[\frac{Z(N,c_j)}{N} - Q\frac{\log N}{N}\right], \tag{32}$$

Thereby, the position and number of active taps are determined by minimizing the RHS of (32), which is equal to maximize

$$Z'(N,\mathbf{c}^k) \triangleq \sum_{j=1}^{k}[Z(N,c_j) - Q\log N] \tag{33}$$

Obviously, $Z'(N, \mathbf{c}^k)$ is a monotonically increasing function of k if $[Z(N, c_j) - Q \log N] > 0$. In this case, for large enough N, active taps will be gotten by the utilization of (33). Herein, $Q = 2{\sigma_r}^2$ [54] is chosen, where ${\sigma_r}^2$ is the power of $r(n)$, and the required threshold is defined as

$$T(N) = \frac{2 \log N}{N} \sum_{n=1}^{N} r^2(n) \tag{34}$$

However, the above criteria are valid only when the white Gaussian noise is used as the input signal. To further expand its application, the cross-correlation relationship should be reduced in $Z(N, p)$ [55], [56]. To achieve this target, the cross-correlation term $S(N, p)$ of $Z(N, p)$ in (25) is analyzed

$$\begin{aligned} &\frac{1}{N} \sum_{n=p+1}^{N} r(n)x(n-p) \\ &\quad = \frac{1}{N} \sum_{n=p+1}^{N} \omega_q(n)x(n-q)x(n-p) \\ &\qquad + \frac{1}{N} \sum_{n=p+1}^{N} \omega_p(n)x(n-p)^2 \\ &\qquad + \frac{1}{N} \sum_{n=p+1}^{N} v(n)x(n-p). \end{aligned} \tag{35}$$

where q and p represent different taps. When $N \to \infty$, $S(N, p) \to \omega_p(n)\sigma_x^2 + \sum_{q \neq p} \omega_q(n) m(q-p)$, and $m(q-p) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=p+1}^{N} x(n-q)x(n-p)$. If the input signal is white Gaussian noise, $m(q-p) = 0$ for all $q \neq p$. Thereby, for enough large N, only tap p has an effect on perceptual activity $Z(N, p)$.

Nevertheless, when the colored noise with a finite correlation length L is taken as the system input, any adjacent active tap q that lags behind the L sample also significantly affects the perceptual activity $Z(N, p)$. In order to reduce the coupling of correlation coefficients, a new activity measurement is proposed and given by

$$\begin{aligned} &\tilde{Z}(N, p) \\ &\quad = \frac{\left[\sum_{n=p+1}^{N} \left(r(n) - y(n) + \hat{\omega}_p(n)x(n-p) \right) x(n-p) \right]^2}{\sum_{n=p+1}^{N} x^2(n-p)}, \end{aligned} \tag{36}$$

which can be used in the case of colored input signals, and then, the cross-correlation term of $\tilde{Z}(N, p)$ can be analyzed in a more intuitive method, and we have

$$\begin{aligned} &\frac{1}{N} \sum_{n=p+1}^{N} \left[r(n) - y(n) + \hat{\omega}_p(n)x(n-p) \right] x(n-p) \\ &\quad = \frac{1}{N} \omega_p \sum_{n=p+1}^{N} x^2(n-p) + \frac{1}{N} \sum_{n=p+1}^{N} v(n)x(n-p) \\ &\qquad + \sum_{q \neq p} \frac{1}{N} \sum_{n=p+1}^{N} [\omega_q - \hat{\omega}_q(n)]x(n-q)x(n-p) \end{aligned} \tag{37}$$

It is observed that the coupling effects can be reduced when the estimated $\hat{w}_q(n)$ converges to w_q in the learning process. Then, the activity threshold is modified to improve the flexibility of the active tap selection, which is given by

$$\tilde{T}(N) = \frac{2 \log(N)}{N} \sum_{n=1}^{N} [r(n) - y(n)]^2. \tag{38}$$

To reduce the influence of time variance of prediction error $\Delta \mathbf{W}(n) = \mathbf{W} - \hat{\mathbf{W}}(n)$ on prediction performance, a forgetting factor is used to the proposed ACD criterion. Then, the positions of the active channel coefficients can be obtained by the following criteria

$$\tilde{Z}^F(n, p) > \tilde{T}^F(n), \tag{39}$$

where

$$\begin{aligned} \tilde{Z}^F(n, p) &= \frac{[\sum_{j=1}^{n} \left(r(j) - y(j) + \hat{\omega}_p(j)x(j-p) \right) x(j-p)F(n, j)]^2}{\sum_{j=1}^{n} x^2(j-p)F(n, j)} \\ \tilde{T}^F(n) &= \frac{2 \log \left[L^F(n) \right]}{L^F(n)} \sum_{j=1}^{n} F(n, j)(r(j) - y(j))^2 \\ F(n, j) &= (\lambda)^{n-j} \quad 0 \ll \lambda < 1 \\ L^F(n) &= \sum_{j=1}^{n} F(n, j). \end{aligned} \tag{40}$$

Then, the ACD scheme is incorporated into the MCC algorithm to construct the ACD-MCC to detect the number and position of active channel coefficients. After that, only the active tap coefficients are updated and the inactive tap coefficients are set to be zero. Based on the ACD scheme and the MCC algorithm, the proposed ACD-MCC algorithm is summarized as follows.

Multiplications per sample interval (MPSI) is used as a measurement to analyse the computational complexity [54]. The classical MCC and NMCC algorithms require $(2M + 2)$ and $(3M + 2)$ MPSIs while the computational complexities for the ACD-MCC algorithm and ACD-NMCC algorithm are $(2M + 2k + 4)$ and $(2M + 3k + 4)$, respectively. It is observed that the proposed ACD-based MCC algorithms have an acceptable computational in comparison with the MCC and NMCC algorithms.

ACD-MCC Algorithm

(1)Initialization

Initialize $S(0,p) = E(0) = L^F(0) = 0$ and $G(n,p) = \delta$, where $0 < \delta \ll {\sigma_u}^2$.

(2)At iteration n,

$$S(n,p) = \lambda S(n-1,p) + \left[r(n) - y(n) + \hat{\omega}_p(n)x(n-p)\right]x(n-p)$$

$$G(n,p) = \lambda G(n-1,p) + x^2(n-p)$$

$$E(n) = \lambda E(n-1) + [r(n) - y(n)]^2$$

$$L^F(n) = \lambda L^F(n-1) + 1$$

$$\tilde{Z}^F(n,p) = \frac{S(n,p)^2}{G(n,p)}$$

$$\tilde{T}^F(n) = \frac{E(n)\log\left[L^F(n)\right]}{L^F(n)}$$

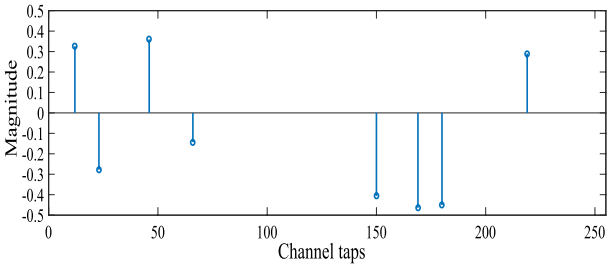
(3)Using the ACD scheme addressed above, the channel coefficient p is detected as an active tap in the case of $\tilde{Z}^F(n,p) > \tilde{T}^F(n)$.

(4)Incorporating the ACD scheme into the MCC, the updating equation of the proposed ACD-MCC algorithm is obtained and given by $\hat{\omega}_p(n+1) = \hat{\omega}_p(n)$

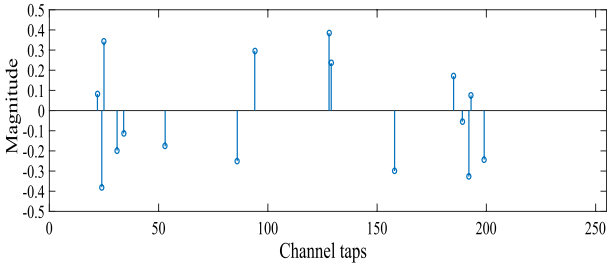
$$+\mu \exp\left(\frac{-e(n)^2}{2\sigma^2}\right)e(n)x(n-p).$$

IV. SIMULATION RESULTS AND DISCUSSIONS

In this paper, the multi-path CIR with length of $M = 256$ and K non-zero coefficients is used to investigate the estimation and tracking performance of the proposed ACD-based MCC algorithms. The typical CIRs of a multi-path channel are presented in Fig. 1. Three different input signals, namely, white Gaussian signal (WGN), colored signal (CN) and speech signal (SS) are used in the simulations. The white Gaussian signal is implemented using a zero-mean Gaussian distribution random signal whose variance is $\sigma_x^2 = 1$, while the colored signal is realized by $x(n) = 0.8x(n-1) + z(n)$, where $z(n)$ is a discrete white zero-mean unit-variance progress. Here, the speech signal has a sampling rate of 8 kHz, which is recorded in our lab. Two kind of impulsive noises are used to verify the estimation performance of the proposed ACD-based algorithms. The normalized misalignment (NM) with the definition of $10\log_{10}\left(\left\|\mathbf{W} - \hat{\mathbf{W}}(n)\right\|_2^2 / \|\mathbf{W}\|_2^2\right)$ is used as a metric to evaluate the behavior of the developed ACD-MCC and ACD-NMCC algorithms. Also, the results are compared with the MCC, NMCC, ZA-MCC, ZA-NMCC, RZA-MCC, RZA-NMCC algorithms. In all the experiments, μ is the step-size while ρ represents the ZA parameters for the ZA-MCC, ZA-NMCC, RZA-MCC and RZA-NMCC algorithms. In addition, $\sigma = 4$ is set in the MCC-type algorithms, while the parameter δ in the RZA-MCC and RZA-NMCC is 10, and the parameter λ in the ACD-based algorithms is set to be 0.99. All simulation lines are obtained from 100 average Monte Carlo runs.



(a) Channel 1 : A Channel with length of 256 and 8 active taps.



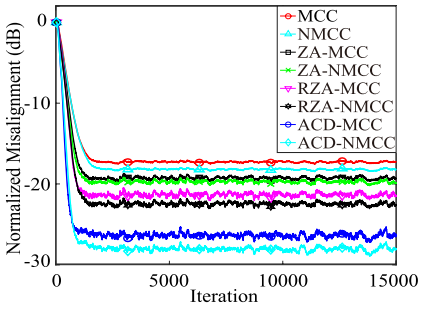
(b) Channel 2 : A Channel with length of 256 and 16 active taps.

FIGURE 1. Two different-sparsity CIRs employed in simulations. (a) Channel 1 : A Channel with length of 256 and 8 active taps. (b) Channel 2 : A Channel with length of 256 and 16 active taps.

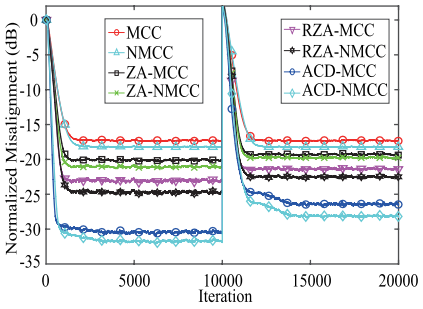
A. THE PERFORMANCE OF THE PROPOSED ACD-BASED MCC ALGORITHMS IN BERNOULLI-GAUSSIAN MIXED NOISE ENVIRONMENT

The discrete-time CIRs are presented in Fig. 1. It is noted that there are only $K = 8$ active channel coefficients and 248 inactive taps in Channel 1, where the Channel 2 consists of $K = 16$ non-zero channel coefficients and 240 inactive ones. Channel 2 is used in Simulation A, while for Simulation B, Channel 1 corresponds to the CIR for the first half of the simulation, and Channel 2 is used in the second half of the simulation. The additive noise $v(n)$ is the sum of Gaussian white noise $s(n)$ and impulse noise $i(n)$, where $s(n)$ is a discrete white zero-mean unit-variance Gaussian noise with a signal-to-noise ratio of 30 dB while the impulsive noise $i(n)$ with a signal-to-interference ratio of 15 dB is generated by $i(n) = b(n)H(n)$, where $b(n)$ is a Bernoulli process whose probability equals to $P[b(n) = 1] = 0.1$.

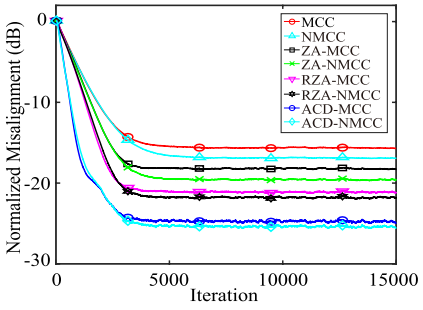
The parameters for the mentioned algorithms in all experiments are conducted under Bernoulli noise distribution environment are shown in Table 1, and the performance analysis for channel 2 under Bernoulli-Gaussian noise environment are demonstrated in Fig. 2. It is found that the proposed ACD-based algorithms achieve the fastest convergence and lowest steady-state NM compared with the mentioned sparse MCC algorithms. Moreover, we can intuitively observe from Fig. 2 that the ACD-NMCC is superior to the ACD-MCC because of the normalization operation. From Simulation B, the proposed ACD-based algorithms shows relatively faster tracking ability than the other algorithms. Also, we can clearly see that with the increase of K, the NMs of ACD-based algorithms are getting worse since the sparsity of the channel is changing to be dense. However, their performance is still better than other algorithms.



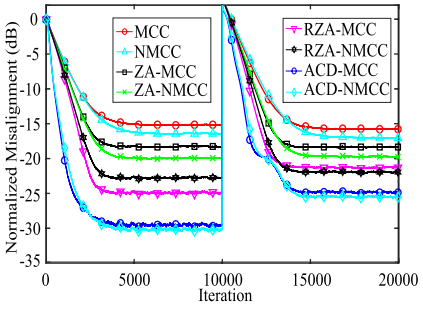
Simulation A: (a) Channel 2 with WGN input



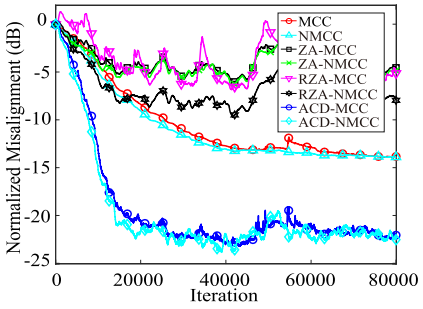
Simulation B: (a) WGN input



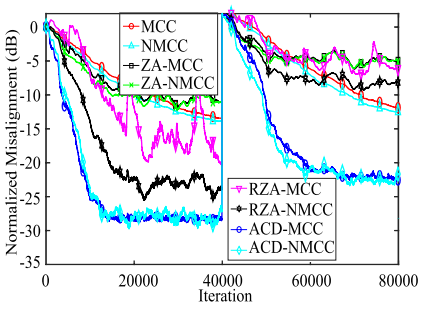
Simulation A: (b) Channel 2 with CN input



Simulation B: (b) CN input



Simulation A: (c) Channel 2 with SS input



Simulation B: (c) SS input

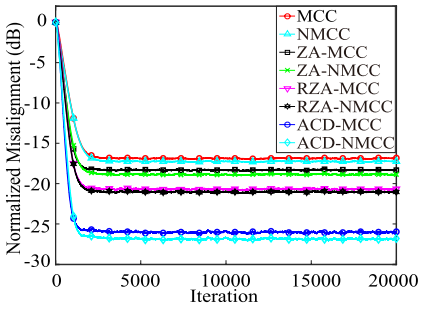
FIGURE 2. **Simulation A, B: Performance of proposed ACD-based algorithms under Bernoulli Gaussian mixed noise environment.**

TABLE 1. **Simulation parameters under Bernoulli Gaussian mixed noise environment.**

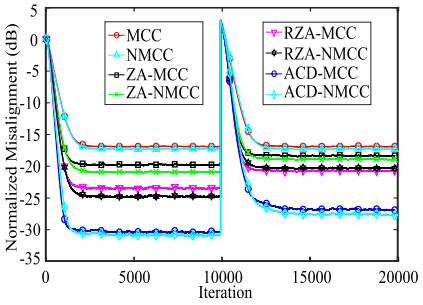
Algorithms \ Input signal	WGN	CN	SS
MCC (μ)	3×10^{-3}	4×10^{-3}	2×10^{-3}
NMCC(μ)	4.2×10^{-2}	5.5×10^{-2}	3.5×10^{-2}
ZA-MCC (μ)	4×10^{-3}	4.5×10^{-3}	4×10^{-3}
ZA-MCC (ρ)	4×10^{-5}	1.3×10^{-5}	4×10^{-5}
ZA-NMCC (μ)	6.3×10^{-2}	6.3×10^{-2}	6.7×10^{-2}
ZA-NMCC (ρ)	5×10^{-5}	1.3×10^{-5}	5×10^{-5}
RZA-MCC (μ)	4.5×10^{-3}	4.5×10^{-3}	5×10^{-3}
RZA-MCC (ρ)	1×10^{-4}	6×10^{-5}	1×10^{-4}
RZA-NMCC (μ)	6×10^{-2}	6.8×10^{-2}	7×10^{-2}
RZA-NMCC (ρ)	9×10^{-5}	4×10^{-5}	9×10^{-5}
ACD-MCC (μ)	1×10^{-2}	1.3×10^{-2}	1.55×10^{-2}
ACD-NMCC (μ)	1.07×10^{-1}	1.7×10^{-1}	1.95×10^{-1}

B. THE PERFORMANCE OF THE PROPOSED ACD-BASED MCC ALGORITHMS IN STUDENT.T DISTRIBUTION NOISE ENVIRONMENT

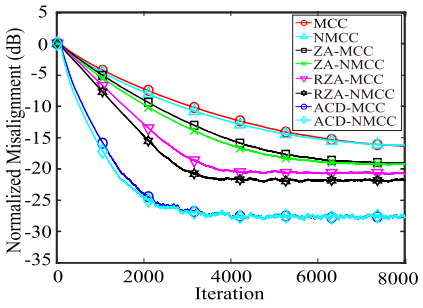
Another non-Gaussian noise environment using Student. T distribution is considered to investigate the performance of the proposed ACD-based algorithms since it has strong tail, and the simulated results are discussed in Fig. 3. Herein, the disturbance noise $v(n)$ is modeled by Student-t distribution with freedom degree of 4. Some parameters are changed to be $\mu_{ACD-NMCC} = 0.16$ for simulation C. Other parameters are presented in Table 2. It can be seen from the simulation C (a) that the ACD-based MCC algorithms provide faster convergence rate and achieves better steady-state performance than those of the MCC-type algorithms when the white Gaussian input signal is used as the input signal. The ACD-MCC algorithm provides about 7 dB gain compared to that of the RZA-MCC algorithm. In contrast, when the input signal are colored signal or speech signal, we can observe that the ACD-based MCC algorithms are superior to the other MCC algorithms, and the performance of the devised ACD-based algorithms are better than the mentioned algorithms. In addition, we can see from the simulation D that the proposed ACD-based algorithms can track the time-varying channel well.



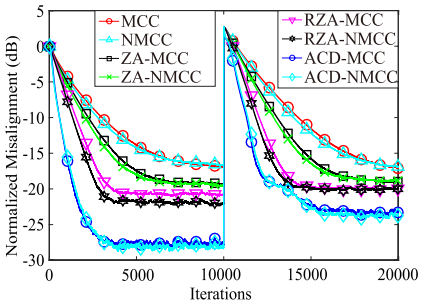
Simulation C: (a) Channel 2 with WGN input



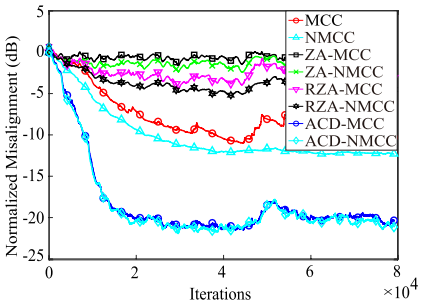
Simulation D: (a) WGN input



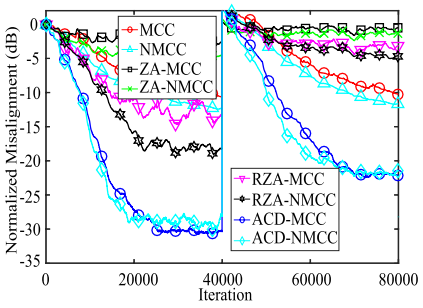
Simulation C: (b) Channel 2 with CN input



Simulation D: (b) CN input



Simulation C: (c) Channel 2 with SS input



Simulation D: (c) SS input

FIGURE 3. Simulation C, D: Performance of proposed ACD-based MCC algorithms under Student-t distribution noise environment.

TABLE 2. Simulation parameters under Student-t distribution noise environment.

Algorithms \ Input signal	WGN	CN	SS
MCC (μ)	2×10^{-3}	2×10^{-3}	3×10^{-3}
NMCC(μ)	3×10^{-2}	3.4×10^{-2}	4×10^{-2}
ZA-MCC (μ)	3×10^{-3}	2.5×10^{-3}	3×10^{-3}
ZA-MCC (ρ)	1.5×10^{-4}	8×10^{-6}	1.5×10^{-4}
ZA-NMCC (μ)	4×10^{-2}	4.2×10^{-2}	4×10^{-2}
ZA-NMCC (ρ)	1×10^{-4}	1×10^{-5}	1×10^{-4}
RZA-MCC (μ)	3×10^{-3}	3.5×10^{-3}	3.5×10^{-3}
RZA-MCC (ρ)	5×10^{-4}	3×10^{-5}	2×10^{-4}
RZA-NMCC (μ)	4×10^{-2}	6×10^{-2}	2×10^{-2}
RZA-NMCC (ρ)	4×10^{-4}	4.5×10^{-5}	1×10^{-4}
ACD-MCC (μ)	4.1×10^{-3}	1×10^{-2}	1×10^{-2}
ACD-NMCC (μ)	5.4×10^{-2}	1.4×10^{-1}	1×10^{-1}

C. CONVERGENCE OF THE ACD-BASED MCC ALGORITHMS

The convergence experiment for the ACD-based algorithms is carried out in Bernoulli mixed noise environment, where the colored signal is used as the input signal. In addition, the l_0-NMCC algorithm which is inspired by [27], [28] is used as a comparison algorithm to verify the superiority of the ACD-based algorithms. In order to ensure that all the algorithms have the same NM level, the parameters are modified to be $\mu_{\text{MCC}}0.003$, $\mu_{\text{NMCC}}\mu_{\text{ACD-MCC}}0.05$, $\mu_{\text{ZA-MCC}}$ 0.0045, $\mu_{\text{ZA - NMCC}}\mu_{\text{RZA - NMCC}}0.0083$, $\mu_{\text{RZA-MCC}}0.006$, $\mu_{\text{ACD-NMCC}}0.6$, $\mu_{l_0\text{ - NMCC}}0.09$, $\rho_{\text{ZA - MCC}} = 1 \times 10^{-5}$, $\rho_{\text{ZA - NMCC}} = 2 \times 10^{-5}$, $\rho_{\text{RZA - MCC}} = 5 \times 10^{-5}$, $\rho_{\text{RZA - NMCC}} = 1 \times 10^{-4}$, $\rho_{l0-\text{NMCC}} = 1.5 \times 10^{-5}$. The obtained simulation results are presented in Fig. 4. It can be seen that the proposed ACD-based MCC algorithms can provide the fastest convergence speed in comparison with the mentioned ACE algorithms. Especially, the proposed ACD-NMCC has the fastest convergence and lowest NM since the ACD scheme can precisely seek the position of the active channel coefficients and ignore the updating of the inactive taps during the learning procedure.

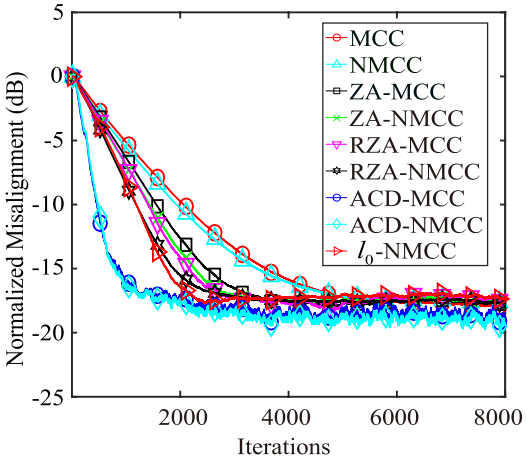


FIGURE 4. Convergence of the ACD-based MCC algorithms.

D. INFLUENCE OF SNR ON THE ACD-BASED MCC ALGORITHMS

The steady-state NM performance of the proposed ACD-based algorithms are investigated in different signal-to-noise ratios (SNRs). The colored signal and Bernoulli distributed noise environment are considered in this experiment. The experimental parameters are the same as that in the Table 1, and the simulation results with different SNRs are presented in Fig. 5. It is noted that the NM is reduced when SNR ranges from 5 dB to 40 dB. Also, the proposed ACD-based MCC algorithms achieve the lowest NM for any SNRs.

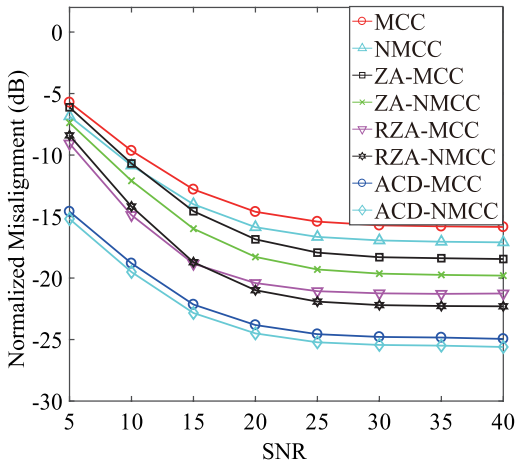


FIGURE 5. Effects of SNRs on the ACD-based MCC algorithms.

V. CONCLUSIONS

In this paper, an active coefficient detection (ACD) criterion is presented and incorporated into the maximum correntropy criterion (MCC)-based adaptive filters to improve the performance for identifying the sparse channels. The proposed ACD-based algorithms have been derived in detail, and their performance has been studied under the different noise environments and different sparsity with three common input signals. The performance obtained from the various simulations showed that the proposed ACD-based MCC algorithms outperform the MCC, NMCC, ZA-MCC, ZA-NMCC, RZA-MCC, RZA-NMCC for dealing with sparse channel estimations.

REFERENCES

[1] H. Huang, W. Xia, J. Xiong, J. Yang, G. Zheng, and X. Zhu, "Unsupervised learning-based fast beamforming design for downlink MIMO," *IEEE Access*, vol. 7, pp. 7599–7605, 2019.

[2] W. Liu, P. P. Pokharel, and J. C. Principe, "The Kernel least-mean-square algorithm," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 543–554, Feb. 2008.

[3] S. Peng, Z. Wu, W. Ma, and B. Chen, "Kernel least mean square based on conjugate gradient," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, New Orleans, LA, USA, Mar. 2017, pp. 2796–2800.

[4] L. Shi and H. Zhao, "Diffusion leaky zero attracting least mean square algorithm and its performance analysis," *IEEE Access*, vol. 6, pp. 56911–56923, Sep. 2018.

[5] G. Gui, L. Xu, and F. Adachi, "Extra gain: Improved sparse channel estimation using reweighted ℓ_1-norm penalized LMS/F algorithm," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Shanghai, China, Oct. 2014, pp. 370–374. doi: 10.1109/ICCChina.2014.7008304.

[6] Y. Li, Z. Jiang, O. M. Osman, X. Han, and J. Yin, "Mixed norm constrained sparse APA algorithm for satellite and network echo channel estimation," *IEEE Access*, vol. 6, pp. 65901–65908, 2018.

[7] W. Shi, Y. Li, L. Zhao, and X. Liu, "Controllable sparse antenna array for adaptive beamforming," *IEEE Access*, vol. 7, pp. 6412–6423, Jan. 2019.

[8] Y. Li, C. Zhang, and S. Wang, "Low-complexity non-uniform penalized affine projection algorithm for sparse system identification," *Circuits, Syst., Signal Process.*, vol. 35, no. 5, pp. 1611–1624, 2016.

[9] Q. Wu, Y. Li, Z. Jiang, and Y. Zhang, "A novel hybrid Kernel adaptive filtering algorithm for nonlinear channel equalization," *IEEE Access*, vol. 7, pp. 62107–62114, 2019.

[10] G. Gui, A. Mehbodniya, and F. Adachi, "Adaptive sparse channel estimation using re-weighted zero-attracting normalized least mean fourth," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Xi'an, China, Aug. 2013, pp. 368–373.

[11] Y. Li, Y. Wang, and T. Jiang, "Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation," *Signal Process.*, vol. 128, pp. 243–251, Nov. 2016.

[12] Y. Li, Y. Wang, and T. Jiang, "Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo cancelation," *AEU Int. J. Electron. Commun.*, vol. 70, no. 7, pp. 895–902, 2016.

[13] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.

[14] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-learning-based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027–3032, Mar. 2019.

[15] J. Wang, J. Yang, J. Xiong, H. Sari, and G. Gui, "SHAFA: Sparse hybrid adaptive filtering algorithm to estimate channels in various SNR environments," *IET Commun.*, vol. 12, no. 16, pp. 1963–1967, Sep. 2018.

[16] E. H. Krishna, K. Sivani, and K. A. Reddy, "OFDM channel estimation using novel LMS adaptive algorithm," in *Proc. Int. Conf. Comput., Commun. Signal Process. (ICCCSP)*, Chennai, India, Jan. 2017, pp. 1–5.

[17] S. Colieri, M. Ergen, A. Puri, and A. Bahai, "A study of channel estimation in OFDM systems," in *Proc. IEEE 56th Veh. Technol. Conf.*, Vancouver, BC, Canada, Sep. 2002, pp. 894–898.

[18] H. Cheng, Y. Xia, Y. Huang, L. Yang, and D. P. Mandic, "A normalized complex LMS based blind I/Q imbalance compensator for GFDM receivers and its full second-order performance analysis," *IEEE Trans. Signal Process.*, vol. 66, no. 17, pp. 4701–4712, Sep. 2018.

[19] Z. Li, Y. Xia, W. Pei, K. Wang, and D. P. Mandic, "An augmented nonlinear LMS for digital self-interference cancellation in full-duplex direct-conversion transceivers," *IEEE Trans. Signal Process.*, vol. 66, no. 15, pp. 4065–4078, Jun. 2018.
[20] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 508–518, Sep. 2000.
[21] Y. Li, Z. Jiang, Z. Jin, X. Han, and J. Yin, "Cluster-sparse proportionate NLMS algorithm with the hybrid norm constraint," *IEEE Access*, vol. 6, pp. 47794–47803, 2018.
[22] Z. Jin, Y. Li, and Y. Wang, "An enhanced set-membership PNLMS algorithm with a correntropy induced metric constraint for acoustic channel estimation," *Entropy*, vol. 19, no. 6, p. 281, Jun. 2017.
[23] W. Ma, D. Zheng, X. Tong, Z. Zhang, and B. Chen, "Proportionate NLMs with unbiasedness criterion for sparse system identification in the presence of input and output noises," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 11, pp. 1808–1812, Nov. 2018.
[24] H. Deng and M. Doroslovacki, "Improving convergence of the PNLMS algorithm for sparse impulse response identification," *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 181–184, Mar. 2005.
[25] R. Baraniuk, "Compressive sensing," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 1–9, Jul. 2007.
[26] Y. Chen, Y. Gu, and A. O. Hero, "Sparse LMS for system identification," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Taipei, Taiwan, Apr. 2009, pp. 3125–3128.
[27] C. Ye, G. Gui, and L. Xu, "Compressive sensing signal reconstruction using L0-norm normalized least mean fourth algorithms," *Circuits, Syst., Signal Process.*, vol. 37, no. 4, pp. 1724–1752, Apr. 2018.
[28] Y. Gu, J. Jin, and S. Mei, "Norm constraint LMS algorithm for sparse system identification," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 774–777, Sep. 2009.
[29] S. Wang, W. Wang, K. Xiong, H. H. C. Iu, and C. K. Tse, "Logarithmic hyperbolic cosine adaptive filter and its performance analysis," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published. doi: 10.1109/TSMC.2019.2915663.
[30] W. Ma, H. Qu, G. Gui, L. Xu, J. Zhao, and B. Chen, "Maximum correntropy criterion based sparse adaptive filtering algorithms for robust channel estimation under non-Gaussian environments," *J. Franklin Inst.*, vol. 352, no. 7, pp. 2708–2727, Jul. 2015.
[31] B. S. Dees, Y. Xia, S. C. Douglas, and D. P. Mandic, "Correntropy-based adaptive filtering of noncircular complex data," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Calgary, AB, Canada, Apr. 2018, pp. 4339–4343.
[32] B. Chen and J. C. Principe, "Maximum correntropy estimation is a smoothed MAP estimation," *IEEE Signal Process. Lett.*, vol. 19, no. 8, pp. 491–494, Aug. 2012.
[33] B. Chen, J. Wang, H. Zhao, N. Zheng, and J. C. Príncipe, "Convergence of a fixed-point algorithm under maximum correntropy criterion," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1723–1727, Oct. 2015.
[34] Y. Li and Y. Wang, "Sparse SM-NLMS algorithm based on correntropy criterion," *IET Electron. Lett.*, vol. 52, no. 17, pp. 1461–1463, Aug. 2016.
[35] L. Shi and Y. Lin, "Convex combination of adaptive filters under the maximum correntropy criterion in impulsive interference," *IEEE Signal Process. Lett.*, vol. 21, no. 11, pp. 1385–1388, Nov. 2014.
[36] W. Shi, Y. Li, and Y. Wang, "Noise-free maximum correntropy criterion algorithm in non-Gaussian environment," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, to be published. doi: 10.1109/TCSII.2019.2914511.
[37] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286–5298, Nov. 2007.
[38] A. Garde, L. Sornmo, R. Jane, and B. F. Giraldo, "Correntropy-based spectral characterization of respiratory patterns in patients with chronic heart failure," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 1964–1972, Aug. 2010.
[39] M. Hajiabadi, B. Razeghi, and M. Mir, "A novel adaptive algorithm for estimation of sparse parameters in non-Gaussian noise," in *Proc. Int. Conf. Workshop Comput. Commun. (IEMCON)*, Vancouver, BC, Canada, Oct. 2015, pp. 1–5. doi: 10.1109/IEMCON.2015.7344492.
[40] Y. He, F. Wang, S. Wang, P. Ren, and B. Chen, "Maximum total correntropy diffusion adaptation over networks with noisy links," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 2, pp. 307–311, Feb. 2019.
[41] B. Chen, L. Xing, J. Liang, N. Zheng, and J. C. Principe, "Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion," *IEEE Signal Process. Lett.*, vol. 21, no. 7, pp. 880–884, Jul. 2014.
[42] S. Wang, L. Dang, B. Chen, S. Duan, L. Wang, and C. K. Tse, "Random Fourier filters under maximum correntropy criterion," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 10, pp. 3390–3403, Oct. 2018.
[43] Y. Li, Z. Jiang, W. Shi, X. Han, and B. Chen, "Blocked maximum correntropy criterion algorithm for cluster-sparse system identifications," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, to be published. doi: 10.1109/TCSII.2019.2891654.
[44] V. C. Gogineni, R. L. Das, and M. Chakraborty, "Proportionate-type hard thresholding adaptive filter for sparse system identification," in *Proc. Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA)*, Siem Reap, Cambodia, Dec. 2014, pp. 1–6. doi: 10.1109/APSIPA.2014.7041807.
[45] W. Ma, B. Chen, J. Duan, and H. Zhao, "Diffusion maximum correntropy criterion algorithms for robust distributed estimation," *Digit. Signal Process.*, vol. 58, pp. 10–19, Nov. 2016.
[46] A. Singh and J. C. Principe, "Using correntropy as a cost function in linear adaptive filters," in *Proc. Int. Joint Conf. Neural Netw.*, Atlanta, GA, USA, Jun. 2009, pp. 2950–2955.
[47] X. Liu, H. Qu, J. Zhao, and B. Chen, "Extended Kalman filter under maximum correntropy criterion," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Vancouver, BC, Canada, Jul. 2016, pp. 1733–1737.
[48] R. Wang, B. Chen, N. Zheng, and J. C. Principe, "A variable step-size adaptive algorithm under maximum correntropy criterion," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Killarney, Ireland, Jul. 2015, pp. 1–5. doi: 10.1109/IJCNN.2015.7280711.
[49] J. C. Principe, *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*. Berlin, Germany: Springer-Verlag, 2010.
[50] D. B. Haddad, M. R. Petraglia, and A. Petraglia, "A unified approach for sparsity-aware and maximum correntropy adaptive filters," in *Proc. 24th Eur. Signal Process. Conf. (EUSIPCO)*, Budapest, Hungary, Aug./Sep. 2016, pp. 170–174.
[51] Y. Wang, Y. Li, J. C. M. Bermudez, and X. Han, "An adaptive combination constrained proportionate normalized maximum correntropy criterion algorithm for sparse channel estimations," *EURASIP J. Adv. Signal Process.*, vol. 2018, pp. 1–13, Sep. 2018. doi: 10.1186/s13634-018-0581-5.
[52] J. Homer and I. Mareels, "LS detection guided NLMS estimation of sparse systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Montreal, QC, Canada, May 2004, p. ii-861.
[53] J. Homer, I. Mareels, R. R. Bitmead, B. Wahlberg, and F. Gustafsson, "LMS estimation via structural detection," *IEEE Trans. Signal Process.*, vol. 46, no. 10, pp. 2651–2663, Oct. 1998.
[54] J. Homer, "Detection guided LMS estimation of sparse channels," in *Proc. IEEE GLOBECOM*, Sydney, NSW, Australia, Nov. 1998, pp. 3704–3709.
[55] J. Homer, I. Mareels, and C. Hoang, "Enhanced detection-guided NLMS estimation of sparse FIR-modeled signal channels," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 8, pp. 1783–1791, Aug. 2006.
[56] J. Homer and I. Mareels, "LS detection guided NLMS estimation of sparse systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Montreal, Que, Canada, May 2004, pp. ii–861.
[57] F. Gustafsson, "Estimation of discrete parameters in linear systems," Ph.D. dissertation, Dept. Elect. Eng., Linkoping Univ., Linköping, Sweden, 1992.

ZEYANG SUN is currently pursuing the B.E. degree in electrical and information engineering from Harbin Engineering University, Harbin, China. His research interest includes adaptive filtering algorithms and its applications.

YINGSONG LI received the B.S. degree in electrical and information engineering and the M.S. degree in electromagnetic field and microwave technology from Harbin Engineering University, in 2006 and 2011, respectively, and the Ph.D. degree from the Kochi University of Technology (KUT), Japan, and Harbin Engineering University, China, in 2014.

He was a Visiting Scholar with the University of California, Davis, from 2016 to 2017. He is a Full Professor with Harbin Engineering University and he is a Visiting Professor with Far Eastern Federal University (FEFU) and KUT. His recent research interests are mainly in remote sensing, underwater communications, signal processing, radar, SAR imaging, compressed sensing and antennas.

Dr. Li is a Senior Member of the Chinese Institute of Electronics (CIE). He is an Associate Editor of the IEEE Access and *Applied Computational Electromagnetics Society Journal*. He is also an Area Editor of *AE Ü-International Journal of Electronics and Communications*. He also serves as a Reviewer for more than 20 journals.

ZHENGXIONG JIANG received the B.S. degree in electrical and information engineering from Harbin Engineering University, Harbin, China, in 2016, where he is currently pursuing the M.E. degree in information and communication engineering.

WANLU SHI was born in Daqing, Heilongjiang, China, in 1992. He received the B.S. degree in information and communication engineering from Harbin Engineering University, in 2015, where he is currently pursuing the M.D.–Ph.D. degree. His project is sparse signal processing. From 2015 to 2016, he was an Assistant Engineer with the 14th Institute of China Electronics Technology Group.