Robust Sparse Channel Estimation Based on Maximum Mixture Correntropy Criterion

Mingfei Lu, Lei Xing, Nanning Zheng, Badong Chen*
Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an,710049, China.
Email: lmfjy8@163.com, xl2010@stu.xjtu.edu.cn, nnzheng@mail.xjtu.edu.cn, chenbd@mail.xjtu.edu.cn.

***Abstract*—The sparse channel estimation problem is drawing increasing attention in broadband wireless communication. Sparsity nature in structure and noise is one of the most important issues in such problems. Researchers have devised several sparsity penalty terms such as zero-attracting (ZA) and correntropy induced metric (CIM) to exploit potential sparse structure information and improve sparse channel estimate accuracy. To combat impulsive (sparse) noises, several adaptive filtering algorithms based on the maximum correntropy criterion (MCC) have been developed, which can achieve excellent performance, especially in heavy-tailed noises. Recently, the concept of mixture correntropy and the maximum mixture correntropy criterion (MMCC) were proposed to further promote the robustness of the MCC against impulsive noises. In this paper, a new robust and sparse adaptive filtering algorithm is developed to estimate the sparse channels with impulsive noises by combining the MMCC and CIM penalty. Thanks to the desirable property of the mixture correntropy, the proposed method behaves quite well with excellent convergence performance. Simulation results show that the new method can outperform several existing methods in sparse channel estimation including the MCC based methods.**

***Index Terms*—sparse channel estimation, adaptive filtering, maximum mixture correntropy criterion, mean square deviation**

I. Introduction

In broadband wireless communication scenarios, signal transmission channel often exhibits sparse structure in time-domain. It is supported by very few dominant coefficients while most of the channel taps are zero [1]. Hence, it takes a very important position to exploit the sparsity in channels and estimate them accurately for realizing reliable wireless communications. Due to their simplicity and easy implementation, adaptive filtering algorithms draw high attention to solve the above sparse channel identification problems. In [2], the proportionate-type algorithms were proposed to estimate the channel parameters, where each channel coefficient is updated in proportion to its estimated magnitude to exploit the sparse information. Several methods were developed to improve the proportionate-type algorithms [3], [4]. By adding different penalty terms to the well-known minimum mean square error (MMSE) criterion, researchers have proposed some effective methods for sparse adaptive filtering, such as zero-attracting LMS (ZALMS) [5], reweighted zero-attracting LMS (RZA-LMS) [5], sparse recursive least squares (RLS) [6] and some other variants [7], [8]. Gui *et al*. [9] proposed a method by incorporating the CIM (correntropy induced metric) into RLS (RLS-CIM), where, with a small kernel width, the CIM provides a good approximation to l_0-norm and can excellently exploit the channel sparsity.

In fact, the MMSE based algorithms provide a good performance under Gaussian noise assumption, but will deteriorate when facing non-Gaussian noises. However, non-Gaussian noises are common in signal transmission channels because of the impulsive property in man-made electromagnetic interference as well as nature noise. As is well known, maximizing the correntropy can effectively suppress impulse noises [10]–[13]. Therefore, it has great potential to introduce MCC to combat the non-Gaussian noise in sparse channels. Ma *et al.* [14] offered a robust sparse adaptive filtering algorithm by combining the MCC with the CIM term called CIM-MCC, which provided a very excellent performance on sparse channel estimation with impulse noises.

Recently, Chen *et al.* [15] proposed the concept of mixture correntropy and extended the MCC to MMCC (maximum mixture correntropy criterion), based on which several learning algorithms (ELMMMCC, KMMC and KRMMC) were developed to improve the performance against complex impulsive noises. Inspired by this, we propose in the present paper a new sparse adaptive channel estimation algorithm based on the MMCC with a CIM penalty, named CIM-MMCC. Specifically, the MMCC is adopted to improve the performance against impulsive noises while the CIM item to exploit the channel sparsity. We theoretically analyzed the mean performance of the proposed algorithm. The convergence performance and parameter sensitivity are studied in simulations under non-Gaussian noises.

The rest of this paper is organized as follows. In Section II, the preliminaries about the correntropy induced metric and the mixture correntropy are revisited. Then, the sparse channel estimation problem is modeled and the new adaptive filtering algorithm CIM-MMCC is developed in Section III. The mean performance analysis of the proposed method is shown in the latter part of this section. Experimental results are presented in Section IV. Finally, Section V gives the conclusion.

This work was supported by National Natural Science Foundation of China (91648208, 61976175), National Natural Science Foundation-Shenzhen Joint Research Program (U1613219), and The Key Project of Natural Science Basic Research Plan in Shaanxi Province of China (2019JZ-05).

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

II. Preliminaries

A. Correntropy and CIM

Given two random vectors X and Y, the correntropy is a nonlinear similarity measure in the kernel space [16]:

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

where $E[\cdot]$ denotes the expectation operator, $\kappa(\cdot,\cdot)$ is a shift-invariant Mercer kernel, and $F_{XY}(x,y)$ represents the joint distribution function. In pratical scenarios, the data distribution is usually unknown and only a finite number of samples $\{x_i, y_i\}(i=1,\ldots,N)$ are available. Thus, the correntropy can be empirically estimated as:

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

The most widely used kernel in correntropy is the Gaussian kernel:

$$\kappa(x,y) = G_\sigma(e) = \exp\left(-\frac{e^2}{2\sigma^2}\right), \quad (3)$$

where $e = x - y$ and $\sigma > 0$ denotes the kernel width.

In this paper, a CIM based sparse penalty is applied to develop an efficient sparse channel estimation algorithm. For a better understanding, we briefly review the CIM below. In a sample space of vector X and Y, the CIM is defined by [17]

$$CIM(X,Y) = \sqrt{G_\sigma(0) - \hat{V}(X,Y)}, \quad (4)$$

Where $G_\sigma(0) = 1$ according to equation (3). It has been proven that if elements $|x_i| > \delta$, where δ is a small positive number, then as $\sigma \to 0$, with CIM we can get arbitrarily close to the l_0-norm of the vector X [17]:

$$\|X\|_0 \sim CIM^2\left(X, \vec{0}\right) = \frac{1}{N}\sum_{i=1}^{N}\left(1 - \exp\left(-\frac{x_i^2}{2\sigma}\right)\right). \quad (5)$$

Consequently, the CIM provides a good approximation to the l_0-norm of a vector and can be adopted as a desirable penalty term to exploit the sparsity structure of a signal tranmission channel.

B. Mixture Correntropy

As a nonlinear similarity measure, correntropy has been successfully used as an efficient optimization cost in signal processing and machine learning. For instance, the regression problem can be solved by maximizing the correntropy $\hat{V}$ (or minimizing the C−Loss: $1 - \hat{V}$) between the model output and the desired response. The MCC has been shown to be very robust against impulsive noises or large outliers [13], [14]. To further improve the performance against complex impulsive noises, a mixture correntropy was defined by [15]

$$\hat{M}(X,Y) = \epsilon\hat{V}_1(X,Y) + (1-\epsilon)\hat{V}_2(X,Y), \quad (6)$$

where $\hat{V}_1$ and $\hat{V}_2$ are two empirical correntropies with different kernels, and ϵ is the mixture coefficient. With the mixture correntropy, researchers proposed the maximum mixture correntropy criterion (MMCC) and adopted it to develope some effective algorithms for adaptive filtering and machine learning. It was shown that adaptive algorithms with MMCC can provide superior performance over those with MCC [15]. In this work, we will utilize the MMCC to combat complex non-Gaussain noises in the scenario of sparse channel estimation.

III. CIM-MMCC Algorithm

A. Problem Description and CIM-MMCC Algorithm

In many wireless communication applications, the signal transmission channel can be modeled as a linear system like a finite impulse response (FIR) filter, in which the observed output signal is:

$$d_n = W_0^T X_n + v_n, \quad (7)$$

where $X_n = [x_{n-M+1}, x_{n-M+2}, \cdots, x_n]^T$ represents the n-th input signal, $W_0 = [\omega_{01}, \omega_{02}, \cdots, \omega_{0M}]^T$ is the channel parameter vector, v_n denotes the additive random noise variable, and M is the size of the channel memory. It is assumed that the channel parameters are real-valued with sparse structure, i.e., most of the channel coefficients will be set to zero. In many pratical systems, the noise v_n is described as non-Gaussian because of the impulsive nature in man-made electromagnetic interference as well as nature noise.

To suppress the impulsive noises and exploit the sparsity of the system as well, a new cost function by combining the MMCC and the CIM is proposed as follows:

$$J_{CIM-MMCC} = -\hat{M}(\hat{y}_n, d_n) + \lambda CIM^2\left(W_n, \vec{0}\right), \quad (8)$$

where the weight factor $\lambda > 0$ is a regularization parameter to balance MMCC term and CIM sparsity term, and $\hat{y}_n = W_n^T X_n$ is the signal output estimate. Let e_n be the n-th instantaneous error:

$$e_n = d_n - \hat{y}_n = d_n - W_n^T X_n, \quad (9)$$

where $W_n = [\omega_{n1}, \omega_{n2}, \cdots, \omega_{nM}]^T$ is the n-th channel estimate. Substituting Eq.(4) and Eq.(6) to Eq.(8), we obtain

$$\begin{aligned} J_{CIM-MMCC} = {} & \frac{\lambda}{M}\sum_{i=1}^{M}\left(1 - \exp\left(-\frac{\omega_{ni}^2}{2\sigma_3^2}\right)\right) \\ & -\epsilon\exp\left(-\frac{e_n^2}{2\sigma_1^2}\right) - (1-\epsilon)\exp\left(-\frac{e_n^2}{2\sigma_2^2}\right) \end{aligned}. \quad (10)$$

Parameters σ_1, σ_2 and σ_3 denote the kernel widths of the MMCC and CIM, respectively.

In Eq.(8), the MMCC term plays a role to deal with impulsive noises while the CIM term contributes to exploit the channel sparsity. Based on the cost function Eq.(10), a gradient-based adaptive filtering algorithm can be derived as follows:

$$\begin{aligned} W_{n+1} &= W_n - \eta\frac{\partial J_{CIM-MMCC}}{\partial W_n} \\ &= (I - \rho g(W_n))W_n + \mu(\epsilon f_1(e_n) + \varphi f_2(e_n))e_n X_n, \end{aligned} \quad (11)$$

where

$$\begin{cases} \mu = \frac{\eta}{\sigma_1^2}, & f_1(e_n) = \exp\left(-\frac{e_n^2}{2\sigma_1^2}\right), \\ \varphi = \frac{\sigma_1^2(1-\epsilon)}{\sigma_2^2}, & f_2(e_n) = \exp\left(-\frac{e_n^2}{2\sigma_2^2}\right), \\ \rho = \frac{\eta\lambda}{M}, & g(W_n) = \operatorname{diag}\left\{\frac{1}{\sigma_3^2}\exp\left(-\frac{\omega_{ni}^2}{2\sigma_3^2}\right)\right\}. \end{cases} \quad (12)$$

We use the parameter μ as the step-size and ρ for weight factor of the CIM. The proposed algorithm is summarized in Algorithm 1.

Algorithm 1 CIM-MMCC algorithm

Initialize: μ, ρ, ϵ, σ_1, σ_2, σ_3, and $W_1 = \vec{0}$.

while $n = 1, 2, \ldots$ **do**
  Input: $X(n)$ and $d(n)$
  Update: $e(n) = d(n) - W_n^T X(n))$
    $f_1(e_n) = \exp(-\frac{e_n^2}{2\sigma_1^2})$
    $f_2(e_n) = \exp(-\frac{e_n^2}{2\sigma_2^2})$
    $g(W_n) = \text{diag}\left\{\frac{1}{\sigma_3^2}\exp\left(-\frac{\omega_{ni}^2}{2\sigma_3^2}\right)\right\}$
  Output: $W_{n+1} = [I - \rho g(W_n)]W_n + \mu[\epsilon f_1(e_n) + \varphi f_2(e_n)]e_n X_n$
end while

B. Performance Analysis

In following, we theoretically analyze the mean convergence performance of the proposed algorithm. First, several assumptions are presented below.

***Assumption 1**:* The input signal $\{X_n\}$ is independent and identically distributed (i.i.d.) with zero-mean Gaussian distribution.

***Assumption 2**:* The noise signal $\{v_n\}$ is i.i.d. with zero-mean and variance σ_v^2, and is independent of $\{X_n\}$.

***Assumption 3**:* The error nonlinearity $f_k(e_n)$ $(k = 1, 2)$ is independent of the input signal $\{X_n\}$.

***Assumption 4**:* The $\{W_n\}$ and $g(W_n)$ are independent of the $\{X_n\}$.

***Assumption 5**:* The expectation $E(f_k(e_\infty))$ $(k = 1, 2)$ is limited.

Remark: *Assumption 1* and *Assumption 2* are commonly used in adaptive filtering [18]. *Assumption 3* is valid when the weight vector W_n lies in the neighborhood of the optimal solution W_0.

The channel estimate weight error vector is

$$\tilde{W}_n = W_0 - W_n. \tag{13}$$

Let the mean of $\tilde{W}_n$ be denoted by $\delta_{\tilde{W}_n} = E\left(\tilde{W}_n\right)$. Substituting Eq.(11) and Eq.(12) to Eq.(13) yields

$$\tilde{W}_{n+1} = A_n\tilde{W}_n + \rho g(W_n)W_0 - \mu[\epsilon f_1(e_n) + \varphi f_2(e_n)]v_n X_n. \tag{14}$$

Taking the expectations on both sides of Eq.(14) and using the independence *Assumptions 1−4*, we obtain

$$E\left(\tilde{W}_{n+1}\right) = E(A_n)E\left(\tilde{W}_n\right) + \rho E(g(W_n))W_0, \tag{15}$$

where A_n is a diagonal matrix and its elements are

$$a_{ii} = 1 - \mu[\epsilon f_1(e_n) + \varphi f_2(e_n)]x_{ni}^2 - \rho g(\omega_{ni}). \tag{16}$$

Apparently, as long as the parameters for the CIM-MMCC are set properly, every diagonal element of A_n satisfies the condition that $0 < a_{ii} < 1$, which guarantees the convergence of the proposed algorithm.

When n approaches $+\infty$, we have:

$$E(\tilde{W}_\infty) = \rho E\left((I - A_\infty)^{-1}\right)E(g(W_\infty))W_0. \tag{17}$$

Then with the diagonal property of A_n we can deduce that

$$E\left(\tilde{W}_\infty\right) = W_0 - diag\left\{\frac{1}{1 + \frac{\rho}{\sigma_x^2}\frac{E(g(\omega_{\infty i}))}{\mu[\epsilon f_1(e_\infty) + \varphi f_2(e_\infty)]}}\right\}W_0, \tag{18}$$

where σ_x^2 denotes the variance of X_n. By comparing Eq.(13) with Eq.(18), it implies that

$$E(W_\infty) = diag\left\{\frac{1}{1 + \frac{\rho}{\sigma_x^2}\frac{E(g(\omega_{i\infty}))}{\mu[\epsilon f_1(e_\infty) + \varphi f_2(e_\infty)]}}\right\}W_0. \tag{19}$$

It's obvious that $E(g(W_\infty))$ is bounded. As a result, $E(W_n)$ will converge to $E(W_\infty)$ as shown in Eq.(19) under *Assumption 5*.

IV. Simulation Studies

To validate the performance of the proposed CIM-MMCC algorithm, we intend to compare it with the other five algorithms, including the LMP (least mean p-power) [19], MCC [13], ZAMCC [10], RZAMCC [10] and CIM-MCC [10]. In this paper, all the results are averaged over 200 independent Monte Carlo runs. The parameter vector of the unknown time-varying channel is set as

$$\omega_{0i} = \begin{cases} \begin{Bmatrix} 0 & i \neq 10 \\ 1 & i = 10 \end{Bmatrix} & n \leq 2000 \\ \begin{Bmatrix} 0 & i\ is\ even \\ 1 & i\ is\ odd \end{Bmatrix} & 2000 < n \leq 3000 \\ \begin{Bmatrix} -1 & i\ is\ even \\ 1 & i\ is\ odd \end{Bmatrix} & 3000 < n \leq 5000 \end{cases}. \tag{20}$$

In Eq.(20), $i = 1, 2, \ldots, M$ and $M = 20$ is the memory size. The total samples are chosen as 5000 for a single simulation test. We define the sparsity degree D_s as:

$$D_s = 1 - \frac{N_{non-zero}}{M} \tag{21}$$

where the $N_{non-zero}$ is the number of the non-zero tap in W_0. Thus, the channel model defined by Eq. (20) has a sparsity of 0.95 during 1000 to 2000 iterations, then D_s changes to 0.5 when the iterations are between 2000 and 3000, and it equals 0 (non-sparsity) after 3000 iterations.

To sufficiently manifest the performance of the CIM-MMCC algorithm against non-Gaussian noises, two noise models, including the Laplacian (finite variance) and the α-stable (infinite variance), are considered in the following Monto Carlo simulations.

Input signals $\{X_n\}$ are generated randomly according to the standard normal distribution $(N(0, 1))$.

A. Channel Estimation under Laplacian Noise

Given mean value ζ and the scale parameter θ, the probability density function (PDF) of the Laplacian noise is defined by

$$f_L(t) = \frac{1}{2\theta}\exp\left(-\frac{|t-\zeta|}{\theta}\right). \tag{22}$$

For our simulaitons, this noise model can be defined as $V_{Laplacian} = V(\zeta, \theta)$.

The estimation performance is evaluated by mean square deviation (MSD) standard which is defined as:

$$MSD(W_n) = \|\tilde{W}_n\|^2 = \|W_0 - W_n\|^2 \tag{23}$$

1) Convergence behavior: We examine the convergence of the proposed algorithm under noise which is modeled as $V_{Laplacian} = V(0.0, 0.5)$. Parameters for all the associated algorithms are set as in Table I. Such a selection of the parameters can ensure that all the six algorithms have almost the same initial convergence rate.

TABLE I
PARAMETERS FOR SIX ALGORITHMS

Parameter category	Parameters for Algorithms					
	μ	σ_1	ρ	σ_2	σ_3	ϵ, δ
LMP($p=1$)	0.02					
MMC	0.02	2.5				
ZAMCC	0.02	2.5	0.002			
RZAMCC	0.02	2.5	0.002			10.0
CIM-MCC	0.018	2.5	0.0005	0.02		
CIM-MMCC	0.02	2.5	0.0005	4.0	0.02	0.7

Remark: σ in algorithms is set as σ_1 if σ_2 and σ_3 are unnecessary.

The average convergence curves in terms of MSD are shown in Fig.1. One can find that the CIM-type adaptive filtering algorithms, including CIM-MCC and CIM-MMCC, track system changes quite well and achieve lower steady-state MSDs in all the three phases of different sparse degree. It is because the CIM term provides a better approximation to l_0-norm function than the ZA and RZA. Besides, the proposed CIM-MMCC algorithm provides even lower steady-state MSDs with almolst the same convergence speed than the CIM-MCC. It means that the mixture correntropy provides us a new approach to promote convergence performance.

2) Steady-state MSD performance: As shown in Fig.2, we consider different θ from 0.1 to 1.1 to produce distinctive noise distributions and study the steady-state MSDs of six algorithms. It is clear that: **a)** The steady-state MSD performances of all the six adaptive filtering algorithms deteriorate with increasing θ. **b)** The CIM-MMCC dilivers the lowest steady-state MSDs under all the noise distributions in this experiment.

B. Channel Estimation under α-Stable Noise

The α-stable model can provide heavy-tailed noises [20], which occur frequently in communication as well as applications in other fields [21]. In this experiment, we will study the performance of the new algorithm against the α-stable noise, which is modeld by a characteristic function below.

$$f_{\alpha S}(t) = \exp\left\{j\delta t - \gamma|t|^{\alpha}\left[1 + j\beta sign(t)\, S(t,\alpha)\right]\right\}, \tag{24}$$

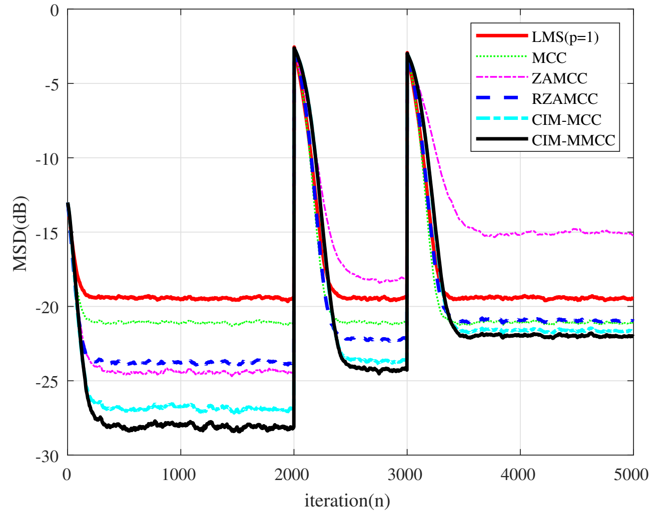

Fig. 1. Convergence behavior of six adaptive filters.

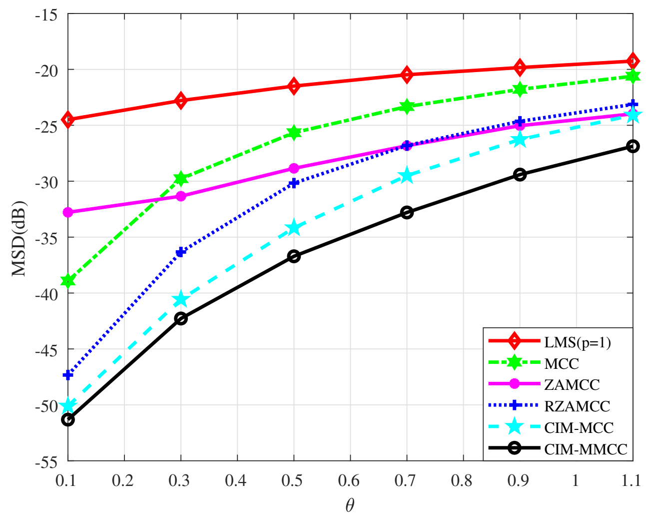

Fig. 2. Steady-state MSDs of six adaptive filters under different θ ($D_s = 0.95$).

in which

$$S(t,\alpha) = \begin{cases} \tan\left(\frac{\alpha\pi}{2}\right) & if \ \ \alpha \neq 1 \\ \frac{2}{\pi}\log|t| & if \ \ \alpha = 1 \end{cases}, \tag{25}$$

where $\alpha \in (0, 2]$ is the characteristic factor, $-\infty < \delta < \infty$ is the location parameter, $\beta \in [-1, 1]$ is the symetric parameter, and $\gamma > 0$ is the dispersion parameter. The characteristic factor α measures the tail heaviness of the distribution. Smaller α makes heavier tail in Eq.(24). In addition, γ measures the dispersion of the distribution. The distribution is symmetric about its location δ when $\beta = 0$. In our simulations, the noise v_n is modeled as $V_{\alpha S} = V(\alpha, \beta, \gamma, \delta)$.

1) Convergence behavior: We consider the problem of sparse channel estimation under impulsive noise model with

$V_{\alpha S} = V(1.2, 0.0, 0.2, 0.0)$. Parameters for all the algorithms are the same as described in Table I. The average convergence curves in terms of MSD are depicted in Fig.3. Convergence performance of the six algorithms are similar to that under the Laplacian noise, and the proposed CIM-MMCC still outperforms other algorithms in all the three phases.

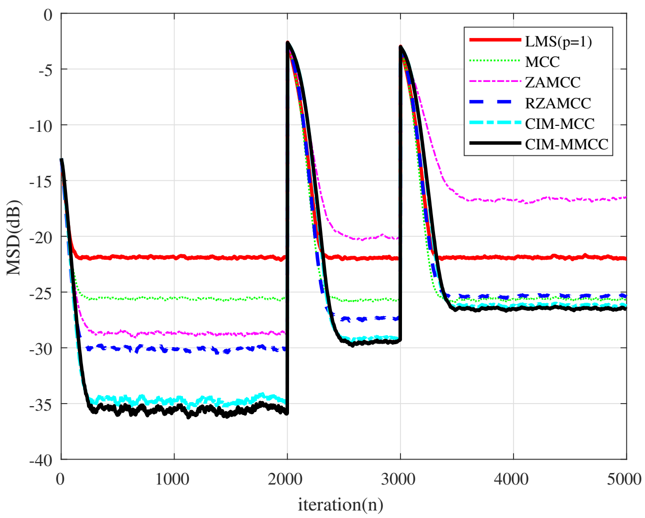

Fig. 3. Convergence performance of six adaptive filters.

2) Steady-state MSD performance: To further evaluate the performance of the proposed algorithm, we conduct the simulations with different values of $\alpha \in [0.5, 1.0, 1.2, 1.5, 1.8]$ and $\gamma \in [0.1, 0.2, 0.5, 1.0, 2.0, 3.0]$. As exhibited in Fig.4 and Fig.5, all the six algorithms perform better with increasing α (implies less heavy tail) and decreasing γ (implies smaller dispersion), and the steady-state MSD curves of our proposed CIM-MMCC are lower than other algorithms with all α and γ alternatives. This suggests that the proposed algorithm shows superior robustness against impulsive noises with different tail-heaviness or dispersion.

C. Parameter sensitivity

The kernel bandwidth σ_1, σ_2 and the mixture coefficient ϵ in MMCC play very important roles in the proposed algorithm. It is necessary to study how the convengence performance will be affected when these parameters are changed. All the following simulations are conducted with the α-stable noise.

First, we investigate the influence of σ_1 in terms of the convergence performance under different distributions of noise. Note that changing σ_1 or σ_2 gives a similar behavior of convergency performance. Alternatives for σ_1 are set as $\sigma_1 \in [0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0]$ and $\sigma_2 = 2.5$ is set as a constant. Steady-state MSD curves are dipicted in Fig.6, from which it can be found that: **a)** The lowest MSDs of the proposed algorithm are obtained when σ_1=1.5 for all the alternatives of α in this example. **b)** The MSD deteriorates as the kernel size increasing from 1.5 to 4.0. **c)** The steady-state MSD performance deteriorates significantly when kernel size

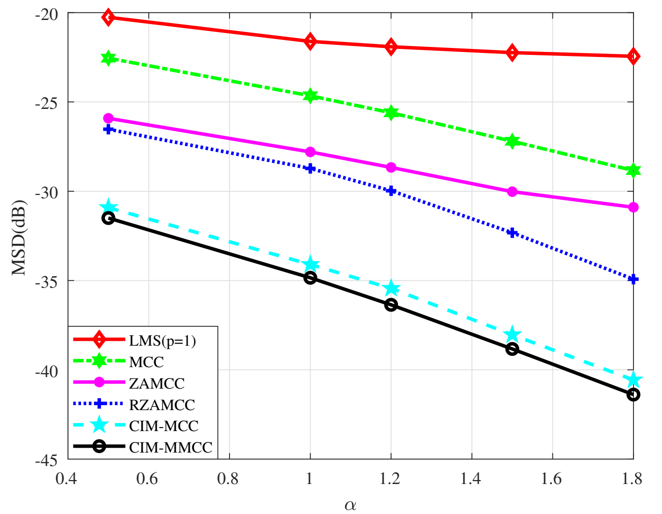

Fig. 4. Steady-state MSDs of six adaptive filters under different α ($D_s = 0.95$).

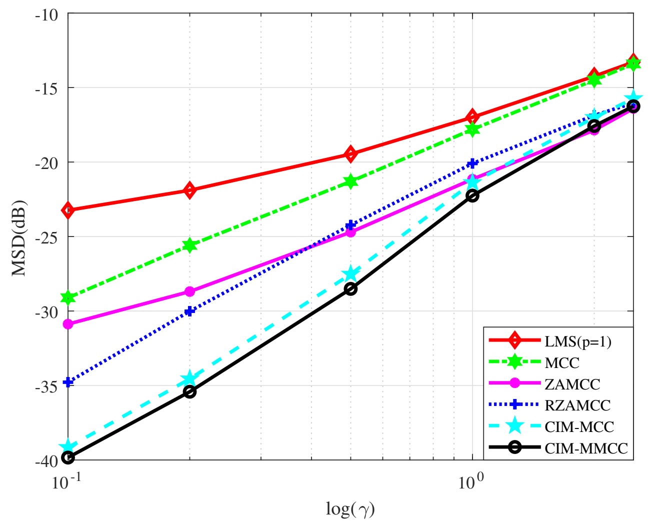

Fig. 5. Steady-state MSDs of six adaptive filters under different γ ($D_s = 0.95$).

σ_1 decreases from 1.5 to 1.0, and then remains unchanged while σ_1 continues to decline. Hence, choosing the suitable kernel size is very important for the CIM-MMCC to achieve desirable MSD performance as well as robustness against different distributions of impulsive noise.

Next, convergence performance of the proposed algorithm is studied when mixture factor ϵ varies inside $[0, 1]$ with different $\sigma_2 \in [1.5, 2.5, 3.0]$ and constant $\sigma_1 = 2.5$. We adjust the step-size μ carefully in each simulation and ensure that all simulations hold approximately the same convergence speed (at the 300-th iteration). From Fig.7, we can conclude that there is no monotonic trend for MSD performance of the CIM-MMCC when ϵ is changing for each σ_2. In this

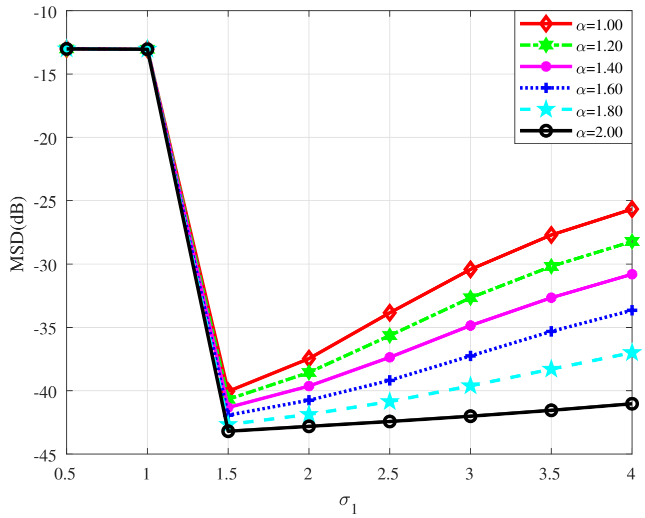

Fig. 6. Steady-state MSDs of CIM-MMCC under different σ_1 and α ($D_s = 0.95$).

experiment, the optimal steady-state MSD is approached when the mixture factor ϵ is chosen between $[0.6, 0.8]$ with σ_2=1.5. This confirms that the CIM-MMCC does have the ability to provide desirable MSD performance with the help of mixture correntropy.

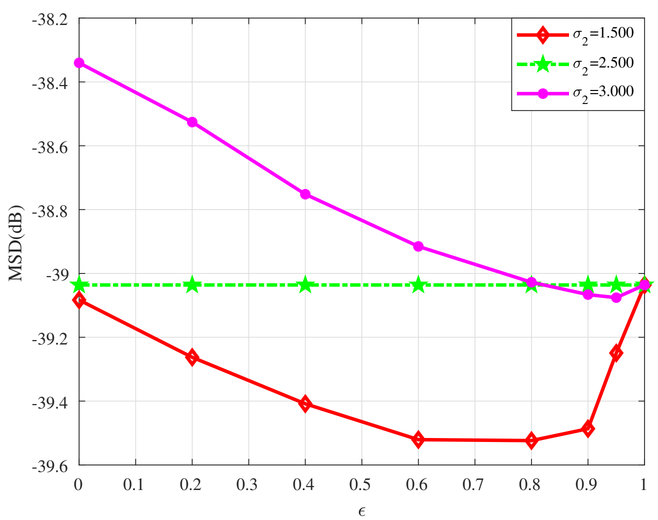

Fig. 7. Steady-state MSDs of CIM-MMCC under different ϵ and σ_2 ($D_s = 0.95$).

V. Conclusion

In this paper, a new robust adaptive algorithm for sparse channel estimation is proposed based on the maximum mixture correntropy criterion (MMCC) combined with the correntropy induced metric (CIM) sparsity penalty term. We theoretically analyse the mean convergence of the proposed algorithm. Simulation results confirm the desirable performance of the new algorithm under both Laplacian (finite variance) and alpha-stable (infinite variance) impulsive noises.

References

[1] L. Davis, V. Hanly, P. Tune, R. Bhaskaran. Channel estimation and user selection in the MIMO broadcast channel, Digit. Signal Process., vol. 21 (5), pp. 608−618, 2011.

[2] D. L. Duttweiler, Proportionate normalized least-mean-squares adaptation in echo cancellers, IEEETrans. Speech, Audio Process., vol. 8 (5), pp. 508−518, 2000.

[3] J. Benesty, S. Gay, An improved PNLMS algorithm, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2002, pp. 1881−1884.

[4] M. Abadi, S. Kadkhodazadeh, A family of proportionate normalized subband adaptive filter algorithms, J. Frankl. Inst., vol. 348 (2), pp. 212−238, 2011.

[5] Y. Chen, Y. Gu, A. Hero, Sparse LMS for system identification, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2009, pp. 3125−3128.

[6] B. Babadi, N. Kalouptsidis, V. Tarokh, SPARLS: the sparse RLS algorithm, IEEE Trans. Signal Process., vol. 8 (8), pp. 4013−4025, 2010.

[7] O. Taheri, S. Vorobyov, Sparse channel estimation with lp-norm and reweighted l1-norm penalized least mean squares, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2011, pp. 2864−2867.

[8] F. Wu, F. Tong, Gradient optimization lp-norm-like constraint LMS algorithm for sparse systeme stimation, Signal Process., vol. 93 (4), pp. 967−971, 2013.

[9] G. Gui, L. Dai, B. Zheng, L. Xu and F. Adachi, Correntropy Induced Metric Penalized Sparse RLS Algorithm to Improve Adaptive System Identification, Vehicular Technology Conference (VTC Spring), Nanjing, 2016, pp. 1-5

[10] S. Zhao, B. Chen and J. C. Prncipe, Kernel adaptive filtering with maximum Correntropy criterion, Proceedings of International Joint Conference on Neural Networks, 2011, pp. 2012−2017.

[11] A. Singh and J. Principe, A loss function for classication based on a robust similarity metric, International Joint Conference on Neural Networks (IJCNN), 2010, pp. 1−6.

[12] B. Chen, L. Xing, J. Liang, N. Zheng and J. C. Prncipe, Steady-state mean square error analysis for adaptive filtering under the maximum correntropy criterion, IEEE Signal Process. Lett., vol. 21 (7), pp. 880−884, 2014.

[13] A. Singh, J. Principe, Using correntropy as cost function in adaptive filters, Proceedings of International Joint Conference on Neural Networks (IJCNN), 2009, pp. 2950−2955.

[14] W. Ma, H. Qu, G. Gui, L. Xu, J. Zhao, B. Chen, Maximum correntropy criterion based sparse adaptive filtering algorithms for robust channel estimation under non-Gaussian environments, J. Franklin Inst., vol. 352 (7), pp. 2708−2727, 2015.

[15] B. Chen, X. Wang, N. Lu, S. Wang, J. Cao, J. Qin, Mixture correntropy for robust learning, Pattern Recognition., vol. 79, pp. 318−327, 2018.

[16] W. Liu, P. Pokharel, J. C. Principe, Correntropy: properties and applications in non-Gaussian signal processing, IEEE Trans. Signal Process., vol. 55 (11), pp. 5286−5298, 2007.

[17] S. Seth, J. Prncipe, Compressed signal reconstruction using the correntropy induced metric, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2008, pp. 3845−3848.

[18] K. Shi, P. Shi, Convergence analysis of sparse LMS algorithms with l1-norm penalty based on white input signal, Signal Process., vol. 90 (12), pp. 3289−3293, 2010.

[19] B. Weng, K. Barner, Nonlinear system identification in impulsive environments, IEEETrans. Signal Process., vol. 53 (7), pp. 2588−2594, 2005.

[20] M. Shao, C. Nikias, Signal processing with fractional lower order moments: stable processes and their applications, Proc. IEEE, vol. 81 (7), pp. 986−1010, 1993.

[21] S. Zhidkov, Impulsive noise suppression in OFDM-based communication systems, IEEETrans. Consum. Electron., vol. 49 (4), pp. 944−948, 2003.