# A Norm Penalized Noise-free Maximum Correntropy Criterion Algorithm

Wanlu Shi\*, Yingsong Li\*,<sup>†</sup> and Felix Albu<sup>‡</sup>

\* College of Information and Communication Engineering, Harbin Engineering University, Harbin, 150001, China

E-mail: shiwanlu@hrbeu.edu.cn

<sup>†</sup> Key Laboratory of Microwave Remote Sensing, National Space Science Center, CAS, Beijing, 100190, China E-mail: liyingsong@ieee.org

<sup>‡</sup> Department of Electronics, Valahia University of Targoviste, Targoviste, 130082, Romania

E-mail: felix.albu@valahia.ro

Abstract-l<sub>1</sub>-norm penalty and noise-free approach are considered in this paper to contribute to a maximum correntropy criterion (MCC) based algorithm. The introduced  $l_1$ -norm constrained noise-free MCC (L<sub>1</sub>-NFMCC) algorithm inherits the good behavior of MCC in non-Gaussian environments. The cost function of the L<sub>1</sub>-NFMCC algorithm is created by introducing  $l_1$ -norm penalty into the traditional cost function of MCC. In this regard, the L<sub>1</sub>-NFMCC algorithm can fully use the sparse characteristics which exist in many real systems. In addition, the noise-free method is used in the L1-NFMCC algorithm to provide a variable convergence step (VCS). The VCS is obtained by minimizing the noise-free (NF) a posteriori error signal with respect to the convergence step. As a consequence, the proposed L<sub>1</sub>-NFMCC algorithm holds an excellent mean square deviation (MSD) behavior. Meanwhile, it shows particularly good property in sparse system. Numerical simulations are utilized to investigate the superiority of the L<sub>1</sub>-NFMCC algorithm in non-Gaussian noises.

#### I. INTRODUCTION

It is no doubt that in adaptive filter (AF) algorithms, selecting a proper cost function is of vital importance [1], [2]. It has previously been observed that the mean square error (MSE) criterion is a feasible choice which performs well in Gaussian environment. The MSE criterion attributes to the developing of least mean square (LMS) algorithm, after which plenty variants of LMS algorithm have been proposed [1], [3], [4], [5], [6], [7]. However, the LMS based AF algorithms cannot deal with non-Gaussian problems well and their behaviors will deteriorate in non-Gaussian environment.

Recent work has been established to use correntropy which is referred as a robust similarity measurement to provide a good choice for constructing a cost function under non-Gaussian environment. Then, the well known maximum correntropy criterion (MCC) algorithm has been established based on correntropy theory [8]. The MCC algorithm has advantages similar with that of the LMS algorithm, meanwhile, it shows superiority in computing impulsive noises [8]. Hence, many scholars focus on the MCC algorithm and lots of variants of MCC have been proposed [8], [9], [10], [11], [12], [13], [14], [15].

The aim of this essay is to explore the sparse MCC based AF algorithm in non-Gaussian environment. It is believed that the non-Gaussian environment can be studied by employing a Gaussian mixture model [16], [17], using which the well known MCC based algorithms have been evaluated [10], [11]. In this regard, the mixture Gaussian environment is employed in this paper to implement an impulsive noise environment.

Recently, a MCC based algorithm using noise-free (NF) approach, namely the NFMCC, has been proposed and investigated in mixture Gaussian environment [18]. The NFMCC algorithm can provide a pretty good behavior in non-Gaussian systems. However, the NFMCC algorithm shows potential to be improved since it didn't consider the a priori information like sparse characteristics.

In this paper, an  $l_1$ -norm constrained NF maximum correntropy criterion (L<sub>1</sub>-NFMCC) algorithm is proposed, which uses the  $l_1$ -norm penalty to take advantage of the sparse nature [19]. Consequently, the L<sub>1</sub>-NFMCC algorithm aims to provide a good performance in non-Gaussian environment, where the cost function of the L<sub>1</sub>-NFMCC algorithm is obtained by considering the norm penalty and the normalized Gaussian kernel. Besides, based on the NF method, a variable convergence step (VCS) is achieved which can make the convergence procedure more fleetly [20], [21]. Several examples are carried out in mixture Gaussian environment to evaluate the L<sub>1</sub>-NFMCC algorithm. After comparing the L<sub>1</sub>-NFMCC algorithm with the MCC, LMS and their variants [18], [22], [23], one can draw a conclusion that the L<sub>1</sub>-NFMCC algorithm has a much better performance in non-Gaussian environment.

# II. The $L_1$ -MCC algorithm

# A. Fundamentals

In Fig. 1, the block diagram of an adaptive filtering system is presented. Generally, the objective of an adaptive filtering algorithm is to get an output signal referred to as  $\mathbf{y}_k$  iteratively which is the closer the better to the ideal output signal  $\mathbf{y}$ , where k is time index.  $\mathbf{w}_k$  is known as the coefficient vector, which is updated iteratively based on the designed adaptive filtering algorithm. In this way, one can get  $y_k = \mathbf{w}_k^T \mathbf{x}_k$ , where the input signal is written as  $\mathbf{x}_k = [x(k), x(k-1), \cdots x(k-N-1)]$ , with N representing the the length of the system.  $d_k$ and  $e_k$  denote, respectively, the desired output signal and the estimation error, while  $n_k$  is the noise with zero mean. Particularly, in this paper, the desired output signal  $d_k$  is decomposed as

$$d_k = d_{NF}(k) + n_k = \mathbf{w}_o^T \mathbf{x}_k + n_k, \qquad (1)$$

where  $d_{NF}(k)$  is referred as the noise-free desired signal and  $\mathbf{w}_o$  stands for the optimal coefficient vector.

Similarly, one can have the following expression for the a priori error signal

$$e_k = d_k - \mathbf{w}_k^T \mathbf{x}_k = d_{NF}(k) - \mathbf{w}_k^T \mathbf{x}_k + n_k = e_{NF}(k) + n_k, \quad (2)$$

where in (2),  $e_{NF}(k)$  is the noise-free a priori error signal.

#### B. The noise-free method

In this paper the noise-free technique, which is also known as a shrinkage method, is utilized [20], [21]. The denoising method performs perfectly in the AF algorithms [24], [25], which focuses on recovering original signal from the observed one using the model below

$$\mathbf{y}_m = \mathbf{x} + \mathbf{n},\tag{3}$$

where  $\mathbf{n}$  denotes a zero-mean independent noisy vector.

The noise-free method is implemented considering the maximum a posteriori probability (MAP) estimation developed in [26], then according to (3) one can obtain

$$\mathbf{f}(\mathbf{z}) = 0.5||\mathbf{y}_m - \mathbf{A}\mathbf{z}||_2^2 + \gamma||\mathbf{z}||_1, \tag{4}$$

where in (4), z stands for a sparse vector, A acts as the dictionary and  $\beta$  is a parameter that adjusts the representation error and sparsity. In this regard, the following relationship is acquired

$$\mathbf{x} = \mathbf{A}\mathbf{z}.$$
 (5)

In this way, the signal  $\mathbf{x}$  is estimated optimally as

$$\hat{\mathbf{x}} = \mathbf{A}\hat{\mathbf{z}}.$$
 (6)

with

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \mathbf{f}(\mathbf{z}).$$
 (7)



Fig. 1. Block diagram of an adaptive filtering system model

#### C. The $L_1$ -MCC algorithm

Referring to [10], the cost function of the MCC algorithm is implemented by employing a normalized Gaussian kernel function and is give as

$$\mathbf{J}_{k} = \frac{1}{\sqrt{2\pi\sigma}} \cdot \frac{1}{N} \sum_{n=k-N+1}^{k} \exp(\frac{-e_{n}^{2}}{2\sigma^{2}}).$$
(8)

In (8),  $\sigma$  is known as kernel width.

Considering the  $l_1$ -norm penalty which can take advantage of the sparse characteristics, the cost function for L<sub>1</sub>-MCC algorithm is available

$$\mathbf{J}_{l1}(k) = \frac{1}{\sqrt{2\pi\sigma}} \cdot \frac{1}{N} \sum_{n=k-N+1}^{k} \exp(\frac{-e_n^2}{2\sigma^2}) + \lambda_{l1} ||\mathbf{w}_k||_{1}.$$
 (9)

Herein, referring to the gradient descent algorithm, the update function for the  $L_1$ -MCC algorithm can be derived from (9)

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \mu \nabla \mathbf{J}_{l1}(k), \tag{10}$$

where  $\nabla$  represents the gradient operator and  $\mu$  stands for the convergence step. Finally, after computing and simplicity, the update function for the L<sub>1</sub>-MCC algorithm is

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \frac{\mu}{\sqrt{2\pi\sigma^3}} \cdot \exp(-\frac{e_k^2}{2\sigma^2}) e_k \mathbf{x}_k - \lambda_{l1} \operatorname{sgn}(\mathbf{w}_k),$$
(11)

with  $sgn(\cdot)$  acting as the sign function.

## III. The proposed $L_1$ -NFMCC algorithm

The proposed  $L_1$ -NFMCC algorithm is derived based on the denoising approach which has been reviewed in section II-B.

Based on (3) and (4), the main idea is to get the noise-free a priori error signal which can be recovered from the estimation error  $e_k$  using the relationship in (2)

$$\hat{e}_{NF}(k) = \arg\min_{e_{NF}(k)} 0.5|e_k - e_{NF}(k)|^2 + \beta|e_{NF}(k)|,$$
(12)

which yields

$$\hat{e}_{NF}(k) = \text{sign}[e_k] \max(|e_k| - \beta, 0).$$
 (13)

In (13),  $\beta$  is chosen according to

$$\beta = \sqrt{q\sigma_n^2},\tag{14}$$

where q is selected within (1,4) and  $\sigma_n^2$  denotes the noise variance.

Till now, the noise-free a priori error signal has already obtained based on (13). Consequently, the remaining problem is to find out the noise-free a posteriori error signal which is defined as

$$\varepsilon_{NF}(k) = d_{NF}(k) - \mathbf{x}_k \mathbf{w}_{k+1}.$$
 (15)

As the a priori error signal is available, a reasonable solution is to consider the straightforward relationship between the noise-free a priori error signal and the noise-free a posteriori error signal, which is given by

$$\varepsilon_{NF}(k) - e_{NF}(k) = \mathbf{x}_k^T(\mathbf{w}_k - \mathbf{w}_{k+1}).$$
(16)

Using the update equation of  $L_1$ -MCC algorithm in (11), (16) becomes

$$\varepsilon_{NF}(k) - e_{NF}(k) =$$

$$\mathbf{x}_{k}^{T} \left[-\frac{\mu_{k}}{\sqrt{2\pi\sigma^{3}}} \cdot \exp(-\frac{e_{k}^{2}}{2\sigma^{2}})e_{k}\mathbf{x}_{k} + \lambda_{l1}\mathrm{sgn}(\mathbf{w}_{k})\right], \qquad (17)$$

where  $\mu_k$  acts as the convergence step which is alterable in this algorithm.

According to (2), (16) can be further written as

$$\varepsilon_{NF}(k) = e_{NF}(k) - \mu_k s[e_{NF}(k) + n_k] \mathbf{x}_k^T \mathbf{x}_k + \lambda_{l1} \mathbf{x}_k^T \operatorname{sgn}(\mathbf{w}_k),$$
(18)

with

$$s = \frac{1}{\sqrt{2\pi\sigma^3}} \cdot \exp(-\frac{e_k^2}{2\sigma^2}). \tag{19}$$

Subsequently, based on (18),  $E[\varepsilon_{NF}^2(k)]$  can be obtained.  $\mu_k$  is available by minimizing  $E[\varepsilon_{NF}^2(k)]$ 

$$\mu_{k} = \frac{E[s]}{E[s^{2}]} \cdot \frac{E[e_{NF}^{2}(k)]E[\mathbf{x}_{k}^{T}\mathbf{x}_{k}] + \lambda_{rl1}E[e_{NF}(k)]E[\mathbf{x}_{k}^{T}\mathbf{x}_{k}\mathbf{x}_{k}^{T}\mathrm{sgn}(\mathbf{w}_{k})]}{E[(\mathbf{x}_{k}^{T}\mathbf{x}_{k})^{2}] \{E[e_{NF}^{2}(k)] + E[n_{k}^{2}]\}}.$$
(20)

The first order Taylor series is utilized to simplify (20), which yields

$$\mu_{k} = \sqrt{2\pi}\sigma^{3} \cdot \frac{E[e_{NF}^{2}(k)]E[\mathbf{x}_{k}^{T}\mathbf{x}_{k}] + \lambda_{rl1}E[e_{NF}(k)]E[\mathbf{x}_{k}^{T}\mathbf{x}_{k}\mathbf{x}_{k}^{T}\mathrm{sgn}(\mathbf{w}_{k})]}{E[(\mathbf{x}_{k}^{T}\mathbf{x}_{k})^{2}] \{E[e_{NF}^{2}(k)] + E[n_{k}^{2}]\}}.$$
(21)

In the proposed L<sub>1</sub>-NFMCC algorithm,  $E[e_{NF}^2(k)]$  is estimated by the time average

$$\sigma_{NF}^2(k) = \alpha \sigma_{NF}^2(k-1) + (1-\alpha)e_{NF}^2(k), \qquad (22)$$

with  $\alpha$  representing the forgetting factor.

Finally, the update function of the L<sub>1</sub>-NFMCC algorithm is

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \frac{\mu_k \delta_{\mathrm{L}_1\mathrm{NFMCC}}}{\sqrt{2\pi}\sigma^3} \cdot \exp(-\frac{e_k^2}{2\sigma^2}) e_k \mathbf{x}_k - \lambda_{l1} \mathrm{sgn}(\mathbf{w}_k).$$
(23)

#### **IV. SIMULATION RESULTS**

In this section, several simulation results are illustrated aiming to show the superiority of the  $L_1$ -NFMCC algorithm. The mixture Gaussian environment is used in the experiments, which is generated by

$$0.95N(0,10^{-4}) + 0.05N(0,10), (24)$$

where in (24), N(u, v) represents the Gaussian distribution with mean u and variance v. In all simulations, the kernel width is 2.

First, the L<sub>1</sub>-NFMCC algorithm is studied in the case of system identification with a white Gaussian input signal. The sparse channel has totally 16 coefficients, where 1 of them is non-zero. The input signal SNR is 30 dB. Simulation parameters are given in detail, they are  $\mu_{\rm LMS} = 0.037$ ,  $\mu_{\rm MCC} = 0.04$ ,  $\mu_{\rm NMCC} = 0.37$ ,  $\mu_{\rm ZAMCC} = 0.037$ ,  $\mu_{\rm NLMS} = 0.4$ ,  $\mu_{\rm PNMCC} =$ 

0.24,  $\delta_{\text{PNMCC}} = 0.01$ ,  $\rho_{\text{PNMCC}} = 0.3$ ,  $\mu_{\text{PNLMS}} = 0.02$ ,  $\delta_{\text{PNLMS}} = 0.01$ ,  $\rho_{\text{PNLMS}} = 0.01$ ,  $\alpha_{\text{L}_1-\text{NFMCC}} = 0.94$ ,  $q_{\text{L}_1-\text{NFMCC}} = 2.4$ ,  $\delta_{\text{L}_1-\text{NFMCC}} = 0.95$ ,  $\beta_{\text{L}_1-\text{NFMCC}} = 4 \times 10^{-6}$ ,  $\alpha_{\text{NFMCC}} = 0.93$ ,  $q_{\text{NFMCC}} = 3.4$ ,  $\delta_{\text{NFMCC}} = 0.89$ . From the result presented in Fig. 2, the L<sub>1</sub>-NFMCC algorithm realizes the lowest MSD level, which is about -48dB.

For the second case, the tracking behavior of the L<sub>1</sub>-NFMCC algorithm is studied at 30dB with a colored noise input signal which is created by using the first order autoregressive (AR1)  $x_k = 0.8x_{k-1} + n_k$ . The parameters are,  $\mu_{\rm MCC} = 7.2 \times 10^{-4}$ ,  $\mu_{\rm LMS} = 7 \times 10^{-4}$ ,  $\mu_{\rm NMCC} = 0.21$ ,  $\mu_{\rm ZAMCC} = 8 \times 10^{-4}$ ,  $\mu_{\rm NLMS} = 0.2$ ,  $\mu_{\rm PNMCC} = 0.16$ ,  $\delta_{\rm PNMCC} = 0.01$ ,  $\rho_{\rm PNLMS} = 0.01$ ,  $\mu_{\rm PNLMS} = 0.15$ ,  $\delta_{\rm PNLMS} = 0.01$ ,  $\rho_{\rm PNLMS} = 0.01$ ,  $\alpha_{\rm NFMCC} = 0.999$ ,  $\delta_{\rm NFMCC} = 0.99$ ,  $q_{\rm NFMCC} = 1$ ,  $\alpha_{\rm L_1-NFMCC} = 0.97$ ,  $q_{\rm L_1-NFMCC} = 1.3$ ,  $\delta_{\rm L_1-NFMCC} = 0.96$ ,  $\beta_{\rm L_1-NFMCC} = 6 \times 10^{-8}$ .

As demonstrated in Fig. 3, the  $L_1$ -NFMCC algorithm provides the best behavior. Under the colored input signal and mixture Gaussian environment, the proposed algorithm performs better by considering the MSD level.

At last, real-life data is utilized to test the proposed L<sub>1</sub>-NFMCC algorithm. The data is given in [27] which is a measured underwater channel obtained in the South China Sea. The input signal SNR is set to be 10 dB, while other parameters are  $\mu_{\rm MCC} = 0.0031$ ,  $\mu_{\rm LMS} = 0.0035$ ,  $\mu_{\rm NMCC} = 0.5$ ,  $\mu_{\rm NLMS} = 0.6$ ,  $\mu_{\rm PNMCC} = 0.4$ ,  $\delta_{\rm PNMCC} = 0.01$ ,  $\rho_{\rm PNMCC} = 0.3$ ,  $\mu_{\rm PNLMS} = 0.002$ ,  $\delta_{\rm PNLMS} = 0.01$ ,  $\rho_{\rm PNLMS} = 0.01$ ,  $\alpha_{\rm NFMCC} = 0.975$ ,  $q_{\rm NFMCC} = 1$ ,  $\delta_{\rm NFMCC} = 0.99$ ,  $\alpha_{\rm L_1-NFMCC} = 0.96$ ,  $q_{\rm L_1-NFMCC} = 1.2$ ,  $\delta_{\rm L_1-NFMCC} = 0.99$ ,  $\beta_{\rm L_1-NFMCC} = 1 \times 10^{-6}$ , respectively. As demonstrated in Fig. 4, the proposed algorithm works effectively in real-life application and performs better than the newly proposed NFMCC algorithm [18].

#### V. CONCLUSIONS

In this paper, an  $l_1$ -norm penalized noise-free maximum correntropy criterion (L<sub>1</sub>-NFMCC) is proposed. Based on the L<sub>1</sub>-MCC algorithm, an alterable step size is obtained by considering a denoising method. In fact, the L<sub>1</sub>-NFMCC algorithm is a general case of the newly developed NFMCC



Fig. 2. Simulation results in Gaussian mixture environment



Fig. 3. Tracking behavior under colored input signal



Fig. 4. Performance in real-data

algorithm which has been verified to perform well in non-Gaussian environment. The proposed L<sub>1</sub>-NFMCC algorithm uses an  $l_1$ -norm penalty meanly to take advantage of the sparse characteristics in sparse system to accelerate the convergence procedure. Investigated in different conditions, for instance, in different input signals and in real-life data, the L<sub>1</sub>-NFMCC algorithm shows good capacity and works successfully. It is reasonable to predict that if a penalty which is better than the  $l_1$ -norm is employed, the new algorithm can reach a much better performance.

# ACKNOWLEDGMENT

Thank you for Dr. Youwen Zhang from Acoustic Science and Technology Laboratory of Harbin Engineering University for providing the measured underwater channel response data to support the practical application in this paper.

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 3072019CFG0801, and in part by the Ph.D. Student Research and Innovation Fund of the Fundamental Research Funds for the Central Universities under Grant 3072019GIP0808.

# REFERENCES

- [1] A. H. Sayed, Fundamentals of Adaptive Filtering, Wiley, NJ, 2003.
- [2] B. Chen, Y. Zhu, J. Hu, J. C. Principe, System Parameter Identification: Information Criteria and Algorithms. Newnes, 2013.
- [3] E. Walach, B. Widrow, "The least mean fourth (LMF) adaptive algorithm and its family," IEEE Trans. Inform. Theory, vol. IT-30, no. 2, pp. 275-283, 1984.

- [4] S. C. Pei, C. C. Tseng, "Least mean p-power error criterion for adaptive FIR filter," IEEE Journal on Selected Areas in Communications, vol. 12, no. 9, pp. 1540-1547, 1994.
- [5] Y. Li, Y. Wang, T. Jiang, "Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation," Signal Processing, vol.128, pp.243-251, Nov. 2016.
- [6] W. Shi, Y. Li, L. Zhao and X. Liu, "Controllable Sparse Antenna Array for Adaptive Beamforming," in IEEE Access, vol. 7, pp. 6412-6423, 2019.
- [7] Y. Li, Y. Wang, T. Jiang, "Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo cancelation," *AEU - International Journal of Electronics and Communications*, vol.70, no.7, pp.859-902, 2016.
- [8] W. Liu, P. P. Pokharel, J. C. Principe, "Correntropy: properties and applications in non-Gaussian signal processing," IEEE Transactions on Signal Processing, vol. 55, no. 11, pp. 5286-5298, 2007.
- [9] B. Chen, X. Liu, H. Zhao, H., et al. "Generalized correntropy for robust adaptive filtering," IEEE Trans. Signal Process., vol. 64, no. 13, pp. 3376-3387, 2016.
- [10] Y. Li, Y. Wang, R. Yang, et al. "A soft parameter function penalized normalized maximum correntropy criterion algorithm for sparse system identification," Entropy, vol. 19, no. 1, Jan., 2017.
- [11] Y. Li, Y. Wang, "Sparse SM-NLMS algorithm based on correntyopy criterion," Electron. Lett., no. 52, pp. 1461-1463, 2016.
- [12] G. Qian, X. Ning, S. Wang, "Mixture complex correntropy for adaptive filter," IEEE Transactions on Circuits and Systems II: Express Briefs, 10.1109/TCSII.2018.2887111.
- [13] Y. Li, Z. Jiang, W. Shi, X. Han, B. D. Chen, "Blocked maximum correntropy criterion algorithm for cluster-sparse system identification," IEEE Transactions on Circuits and Systems II: Express Briefs, 10.1109/TCSI-I.2019.2891654.
- [14] 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," Journal of the Franklin Institute, vol. 352, no. 7, pp. 2708-2727, July 2015.
  [15] B. Chen, J. Wang, H. Zhao, N. Zheng, J. C. Principe, "Convergence of
- [15] B. Chen, J. Wang, H. Zhao, N. Zheng, J. C. Principe, "Convergence of a Fixed-Point Algorithm under Maximum Correntropy Criterion," IEEE Signal Process. Lett., vol. 22, no. 10, pp. 1723-1727, Oct. 2015.
- [16] E. Daeipour and Y. Bar-Shalom, "An interacting multiple model approach for target tracking with glint noise," IEEE Trans. Aerosp. Electron. Syst., vol. 31, no. 2, pp. 706-715, Apr. 1995.
- [17] M. Svensen and C. M. Bishop, "Robust bayesian mixture modeling," Neurocomputing, vol. 64, pp. 235-252, Mar. 2005.
- [18] W. Shi, Y. Li and Y. Wang, "Noise-free Maximum Correntropy Criterion Algorithm in Non-Gaussian Environment," in processing IEEE Transactions on Circuits and Systems II: Express Briefs, 10.1109/TC-SII.2019.2914511.
- [19] Y. Li, Z. Jiang, O. M. O. 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.
- [20] I. Daubechies, M. Defrise, and C. De-Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," Comm. Pure Appl. Math., vol. 57, pp. 1413-1457, 2004.
- [21] M. Zibulevsky and M. Elad, "11-12 optimization in signal and image processing," IEEE Signal Process. Mag., pp. 76-88, May 2010.
- [22] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," IEEE Transactions on Speech and Audio Processing, vol. 8, no. 5, pp. 508-518, Sep. 2000.
- [23] J. Duan, W. Ma, B. Chen, et al. "Robust proportionate adaptive filter based on maximum correntropy criterion for sparse system identification in impulsive noise environments," Signal, Image and Video Processing, vol. 12, no. 1, pp. 117-124, Jan. 2018.
- [24] M. Z. A. Bhotto and A. Antoniou, "A family of shrinkage adaptive filtering algorithms," IEEE Trans. Signal Process., vol. 61, no. 7, pp. 1689-1697, Apr. 2013.
- [25] Y. M. Shi, L. Huang, C. Qian, et al, "Shrinkage linear and widely linear complex-valued least mean squares algorithms for adaptive beamforming," IEEE Trans. Signal Process., vol. 63, no. 1, pp. 119-131, Jan. 2015.
- [26] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," SIAM J. Sci. Comput., vol. 20, no. 1, pp. 33-61, 1998.
- [27] Y. Zhang, J. Li, Y. V. Zakharow, et al, "Deep learning based single carrier communications over time-varying underwater acoustic channel," IEEE Access, vol.7, pp.38420-38430, 2019.