A Robust Group-Sparse Proportionate Affine Projection Algorithm with Maximum Correntropy Criterion for Channel Estimation

Zhengxiong Jiang, Yingsong Li*
College of Information and Communications Engineering,
Harbin Engineering University,
Harbin 150001, China
Email: *liyingsong@ieee.org

Yuriy Zakharov
Department of Electronic Engineering,
University of York,
York YO10 5DD, United Kingdom
Email: yury.zakharov@york.ac.uk

***Abstract*—In many engineering applications, noise often exhibits strongly impulsive characteristics, while the conventional adaptive filtering (AF) algorithms are less robust to the impulsive noise. The AF algorithms based on maximum correntropy criterion (MCC) have been devised to effectively enhance the adaptive estimation performance in impulsive noise environments. In this paper, a robust group-sparse proportionate affine projection (RGS-PAP) algorithm based on MCC is proposed for estimating group-sparse channels which often occur in network echo paths and satellite communications channels. The constructed RGS-PAP algorithm is derived via exerting a mixed $l_{2,1}$ norm constraint of AF weights into the updating equation of the affine projection algorithm with MCC to utilize the group-sparse characteristics. The developed RGS-PAP algorithm is analyzed by setting up various simulation experiments to verify its robustness and effectiveness. Simulation results indicate that the proposed RGS-PAP algorithm provides faster convergence and lower estimation bias compared with other algorithms under various input signals in impulse noise environments.**

***Index Terms*—Channel estimation, maximum correntropy criterion, PAP algorithm, mixed $l_{2,1}$ norm, impulse noise environments**

I. Introduction

Adaptive filtering (AF) has found a wide range of applications in radar and wireless communication systems, and it has been proved to be effective for system identifications [1]–[3]. In the wireless transmission area, AF has been used for channel estimation which is a vital part in modern communication systems [4]–[6]. Thus, it is important to study the adaptive channel estimation for improving the communication performance [4]–[6]. In addition, AF can accurately estimate the channels via acting as a channel estimator.

The most representative AF algorithm is the least mean square (LMS), which is realized by solving the Wiener-Hopf equation based on the steepest descent method [6]. The LMS and its normalized form (normalized LMS, NLMS) have been extensively studied owing to their simple implementation and stable performance [7]. As an extension of the NLMS algorithm, the affine projection (AP) algorithm reuses the input signal, which significantly reduces the gradient noise [8]. Although the AP scheme enhances the convergence in comparison with the NLMS algorithm for colored noise [9], it is also more complicated. In all these algorithms, the mean square error (MSE) has been considered at the cost function to achieve good estimation performance when the noise obeys the Gaussian distribution. However, in many engineering environments, the noise often is impulse-like non-Gaussian. In these cases, both AP and LMS variants might suffer from performance degradation. To address this problem, the maximum correntropy criterion (MCC) has been presented to develop robust AF algorithms [10], [14].

In practice, many unknown channels to be estimated have sparse characteristics, like the channels in underwater and satellite communication systems [15], [16]. However, classical AF algorithms, such as LMS, AP and MCC algorithms [1], [10]–[13], are unable to take the advantage of the channel sparseness to improve the estimation performance [17]. Recently, the proportionate scheme has been proposed for the NLMS, denoted as PNLMS, that utilizes channel sparseness to arrive at a step-size varying over the AF taps [18]. Then, several improved proportionate-type algorithms have been proposed [19]. Additionally, a series of sparse AF algorithms have been presented from the motivation of the compressed sensing theory, such as zero-attracting (ZA) LMS (ZA-LMS) and its reweighted form (RZA-LMS) [20]–[22]. These algorithms cannot effectively handle the group-sparse channels, where active coefficients of the channel impulse response (CIR) are aggregated into one-group (such as network echo channels) and multi-group (like satellite communication channels) [23], [24]. To utilize the group-sparsity properties, several group-sparse AF algorithms have been proposed [16], [25]–[29]. The block-sparse LMS (BS-LMS) algorithm [25] incorporates the mixed $l_{2,0}$-norm into the cost function; it has a superior performance [25] compared to the LMS and its ZA forms. The block-sparse PNLMS (BS-PNLMS) algorithm in [26] employs the mixed $l_{2,1}$-norm to modify the cost function of the PNLMS for exploiting the group -sparsity [26].

In this paper, we use the MCC criterion as the cost function to improve the AP algorithm and derive a robust group-sparse proportionate AP (PAP) (RGS-PAP) algorithm for dealing

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

with the group-sparse channels in impulse noise environments. The proposed RGS-PAP algorithm is investigated and verified in symmetrical alpha-stable (SαS) distribution noise environment. Experimental results show that the proposed RGS-PAP algorithm converges faster and achieves a lower steady-state error in α-stable impulsive noise environments in comparison with other MCC, NLMS and their proportionate versions algorithms.

II. Review of MCC and PAP Algorithms

A. MCC algorithm

In the identification framework, it is assumed that the input signal is $\mathbf{x}(n) = [x(n), x(n-1), x(n-2), \cdots, x(n-L+1)]^T$, and the CIR is $\mathbf{h}(n) = [h_0(n), h_1(n), \cdots, h_{L-1}(n)]^T$. In this case, n is the time index and L is the length of the filter. Then, the reference signal $d(n)$ is

$$d(n) = \mathbf{h}^T(n)\mathbf{x}(n) + r(n), \tag{1}$$

where $r(n)$ is the measurement noise. The estimated error is given by

$$e(n) = d(n) - y(n) = d(n) - \hat{\mathbf{h}}^T(n-1)\mathbf{x}(n), \tag{2}$$

where $y(n)$ is the AF output and $\hat{\mathbf{h}}(n)$ is the CIR estimate. The MCC algorithm is implemented by using the stochastic gradient method to find the maximum value of the correntropy between $d(n)$ and $y(n)$. The updating equation of the basic MCC algorithm is given by [10]

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mu_{\mathrm{MCC}} \exp\left(-\frac{e^2(n)}{2\sigma^2}\right) e(n)\mathbf{x}(n), \tag{3}$$

in which μ_{MCC} represents the overall step size and σ denotes the kernel width.

B. PAP algorithm

The AP algorithm reuses the information from previous instants to speed up the convergence, particularly when the input signal is colored. The input signal of the AP algorithm is expressed as an $L \times P$ matrix

$$\mathbf{X}(n) = [\mathbf{x}(n), \mathbf{x}(n-1), \cdots, \mathbf{x}(n-P+1)], \tag{4}$$

where P denotes the projection order. The reference signal in the AP algorithm is $\mathbf{d}(n) = [d(n), d(n-1), \cdots, d(n-P+1)]^{\mathrm{T}}$. Then, the output is

$$\mathbf{Y}(n) = \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1) \tag{5}$$

with the estimated error

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{Y}(n). \tag{6}$$

Inspired by the famed PNLMS [18], the proportionate-update method has been introduced into the AP to construct the PAP algorithm whose updating equation is

$$\begin{aligned}&\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \\ &\mu_{\mathrm{PAP}}\mathbf{G}(n-1)\mathbf{X}(n)\left[\mathbf{X}^T(n)\mathbf{G}(n-1)\mathbf{X}(n) + \delta\mathbf{I}_{\mathrm{P}}\right]^{-1}\mathbf{e}(n)\end{aligned} \tag{7}$$

where μ_{PAP} acts as the total the step size, $\delta > 0$ represents a small constant, and $\mathbf{I}_{\mathrm{P}}$ is a $P \times P$ identity matrix. $\mathbf{G}(n-1)$ is the step-size control matrix, which is expressed as [18], [19]

$$\mathbf{G}(n-1) = \mathrm{diag}\left\{g_0(n-1), g_1(n-1), \ldots, g_{L-1}(n-1)\right\} \tag{8}$$

where

$$g_l(n-1) = \frac{\varphi_l(n-1)}{\sum\limits_{i=0}^{L-1}\varphi_i(n-1)}, 0 \leq l \leq L-1, \tag{9}$$

and

$$\varphi_l = \max\left\{p\max\left\{q, \left|\hat{h}_0\right|, \left|\hat{h}_1\right|, \cdots, \left|\hat{h}_{L-1}\right|\right\}, \left|\hat{h}_l\right|\right\}, \tag{10}$$

where the parameters p and q are used to prevent the update from stalling [18].

III. The Proposed RGS-PAP Algorithm

The PAP algorithm achieves fast convergence in Gaussian noise environments, but the performance might be degraded under the impulsive noise. To give a resistance to the impulsive noise and make full use of the group-sparsity of the channel, we propose the RGS-PAP algorithm which is developed exerting the mixed $l_{2,1}$-norm into the cost function of the AP-based MCC. According to the basis pursuit approach (BPA) [30], the proposed RGS-PAP aims to solve the following problem

$$\begin{aligned}&\min_{\mathbf{h}'(n)} \|\mathbf{h}'(n)\|_{2,1}, \\ &\text{subject to } \mathbf{e}'(n) = \left[1 - \mu\exp\left(-\frac{\mathbf{e}(n)\cdot\mathbf{e}(n)}{2\sigma^2}\right)\right]\cdot\mathbf{e}(n),\end{aligned} \tag{11}$$

where $\mathbf{e}'(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\mathbf{h}'(n)$, $\mathbf{h}'(n)$ is the correctiveness component [30], and $\mathbf{e}(n)\cdot\mathbf{e}(n)$ is the Hadamard product. According to the Lagrange multiplier method, we obtain

$$\begin{aligned}&J(n) = \|\mathbf{h}'(n)\|_{2,1} \\ &+\boldsymbol{\lambda}\left(\mathbf{d}(n) - \mathbf{X}^T(n)\mathbf{h}'(n) - [1 - \mu\exp(-\tfrac{\mathbf{e}(n)\cdot\mathbf{e}(n)}{2\sigma^2})]\cdot\mathbf{e}(n)\right),\end{aligned} \tag{12}$$

where $\boldsymbol{\lambda} = [\lambda_0, \lambda_1, \cdots, \lambda_{\mathrm{P}-1}]$ and

$$\|\mathbf{h}'(n)\|_{2,1} = \left\|\begin{bmatrix}\left\|\mathbf{h}'_{[1]}(n)\right\|_2 \\ \left\|\mathbf{h}'_{[2]}(n)\right\|_2 \\ \vdots \\ \left\|\mathbf{h}'_{[K]}(n)\right\|_2\end{bmatrix}\right\|_1 = \sum_{i=1}^{K}\left\|\mathbf{h}'_{[i]}(n)\right\|_2, \tag{13}$$

where $K = L/B$, and B is the size of a block. Then, let

$$\frac{\partial J(n)}{\partial\mathbf{h}'(n)} = \mathbf{0}, \frac{\partial J(n)}{\partial\boldsymbol{\lambda}} = \mathbf{0}. \tag{14}$$

We obtain

$$\begin{aligned}&\mathbf{h}'(n) = \mathbf{K}(n)\mathbf{X}(n)\boldsymbol{\lambda}^T, \\ &\mathbf{d}(n) = \mathbf{X}^T(n)\mathbf{h}'(n) + [1 - \mu\exp(-\frac{\mathbf{e}(n)\cdot\mathbf{e}(n)}{2\sigma^2})]\cdot\mathbf{e}(n),\end{aligned} \tag{15}$$

$\mathbf{K}(n)$ is a new step-size control matrix which will be given next. Solving the equation in (15) results in

$$\mathbf{h}'(n) = \mathbf{K}(n)\,\mathbf{X}(n)\left[\mathbf{X}^T(n)\mathbf{K}(n)\,\mathbf{X}(n)\right]^{-1} \cdot \left[\mathbf{d}(n) - [1 - \mu \exp(-\frac{\mathbf{e}(n)\cdot\mathbf{e}(n)}{2\sigma^2})]\cdot \mathbf{e}(n)\right] \tag{16}$$

Observing equation (16), we found that it is difficult to get a solution of the equation (16) [30], and then, we approximate the $\mathbf{K}(n)$ by $\mathbf{K}(n-1)$, so that (15) can be rewritten as

$$\mathbf{h}'(n) = \mathbf{K}(n-1)\,\mathbf{X}(n)\left[\mathbf{X}^T(n)\mathbf{K}(n-1)\,\mathbf{X}(n)\right]^{-1} \cdot \left[\mathbf{d}(n) - [1 - \mu \exp(-\frac{\mathbf{e}(n)\cdot\mathbf{e}(n)}{2\sigma^2})]\cdot \mathbf{e}(n)\right]. \tag{17}$$

Based on the BPA [30], the updating equation of the devised RGS-PAP algorithm is given by

$$\hat{\mathbf{h}}(n) = \mathbf{R}(n)\,\hat{\mathbf{h}}(n-1) + \mathbf{h}'(n), \tag{18}$$

where $\mathbf{R}(n)$ is defined as

$$\mathbf{R}(n) = \mathbf{I} - \mathbf{K}(n-1)\,\mathbf{X}(n)\left[\mathbf{X}^T(n)\mathbf{K}(n-1)\,\mathbf{X}(n)\right]^{-1}\mathbf{X}^T(n), \tag{19}$$

in which $\mathbf{I}$ is an $L \times L$ identity matrix. From equations (17), (18) and (19), we can obtain the final updating equation of the RGS-PAP algorithm:

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mu\mathbf{K}(n-1)\,\mathbf{X}(n) \cdot\left[\mathbf{X}^T(n)\mathbf{K}(n-1)\,\mathbf{X}(n) + \delta\mathbf{I}\right]^{-1} \exp(-\frac{\mathbf{e}(n)\cdot\mathbf{e}(n)}{2\sigma^2})\cdot \mathbf{e}(n), \tag{20}$$

in which $\delta > 0$ is a small constant. The step-size control matrix $\mathbf{K}(n-1)$ is given by

$$\mathbf{K}(n-1) = \operatorname{diag}\left\{m_1(n-1)\,\mathbf{1}_B, m_2(n-1)\,\mathbf{1}_B, \cdots, m_K(n-1)\,\mathbf{1}_B\right\}, \tag{21}$$

where

$$m_k(n-1) = \frac{\psi_k(n-1)}{\sum\limits_{i=1}^{K}\psi_i(n-1)}, 1 \le k \le K,$$
$$\psi_k(n-1) = \max\left\{p\max\left\{q, \left\|\hat{\mathbf{h}}_{[1]}\right\|_2, \left\|\hat{\mathbf{h}}_{[2]}\right\|_2, \cdots, \left\|\hat{\mathbf{h}}_{[K]}\right\|_2\right\}, \left\|\hat{\mathbf{h}}_{[k]}\right\|_2\right\}, \tag{22}$$

and $\mathbf{1}_B$ is a B-length row vector with all ones.

IV. Simulation Results

In this section, simulation experiments are constructed to analyze the performance of the RGS-PAP algorithm in the context of identification scenarios. Since the SαS distribution ($V_{\mathrm{S}\alpha\mathrm{S}} = (\alpha, 0, \gamma, 0)$) can be modeled to create the non-Gaussian phenomenon which is ubiquitous in practice, we choose it to implement the impulsive noise. The signal-to-noise ratio (SNR), defined by $\mathrm{SNR} = 10\log_{10}\frac{\delta_x^2}{\delta_r^2}$, is set to be 20 dB in the following experiments. Herein, δ_x^2 is the power of the input signal, while δ_r^2 is the power of the noise obtained from $V_{\mathrm{S}\alpha\mathrm{S}}$. Two group-sparse channels of length L=1024, one of which is the one-group sparse channel with active coefficients within taps [257,272], and the other is a two-group sparse channel with active coefficients within taps [257,272] and [769,784], are considered to evaluate the developed RGS-PAP algorithm. In all simulation experiments, $\alpha = 1.4$ and $\gamma = 0.2$ are used to construct the impulsive noise. The other parameters are $\sigma = 0.9$, $\delta_{\mathrm{AP}} = \delta_{\mathrm{RZA-AP}} = 0.01$ and $\delta_{\mathrm{RGS-PAP}} = \delta = \delta_{\mathrm{PNMCC}} = \frac{1}{L}\delta_{\mathrm{AP}}$ [31]. Two kind of input signals, namely, white Gaussian and colored signals, are employed to investigate the performance of the RGS-PAP algorithm. The colored signal is obtained from the white Gaussian signal by filtering through a first-order system with a pole at 0.8. The normalized misalignment (NM) is used to evaluate the performance of all algorithms, which is given by $10\lg(\left\|\mathbf{h}-\hat{\mathbf{h}}\right\|_2^2 / \|\mathbf{h}\|_2^2)$.

A. Performance with different size B of CIR blocks

The NM is investigated for different B with the colored input and the two-group channel with $\mu = 0.01$, as shown in Fig.1. It can be seen that the best performance is achieved when B=16 is used for the two-group channel.

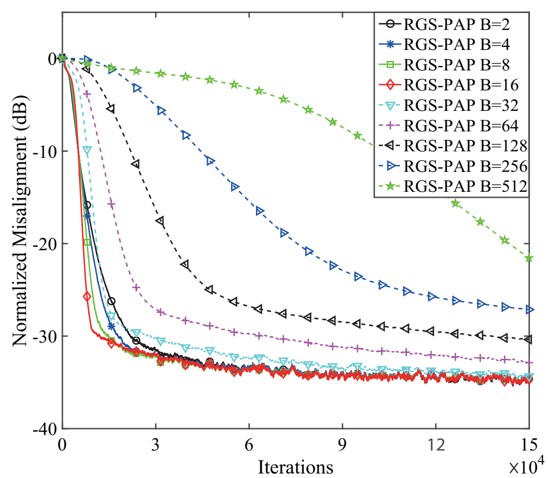

Fig. 1: Effects of B on the RGS-PAP algorithm.

B. Performance with different input signals

In Fig. 2 to Fig. 5, the performance of the RGS-PAP algorithm is further studied and compared with those of the AP, RZA-AP, PAP, MCC, RZA-MCC and PNMCC algorithms. In these experiments, $\mu_{\mathrm{AP}} = 0.02$, $\mu_{\mathrm{RZA-AP}} = 0.06$, $\mu_{\mathrm{PAP}} = 0.01$, $\mu_{\mathrm{MCC}} = 0.0004$, $\mu_{\mathrm{RZA-MCC}} = 0.006$, $\mu_{\mathrm{PAP}} = 0.115$ and $\mu_{\mathrm{RGS-PAP}} = 0.01$ are used. The block size B in the RGS-PAP algorithm is set to B = 16, for which the algorithm provides the best performance. It is clear that the proposed RGS-PAP algorithm achieves the fastest convergence and lowest steady-state NM for the colored input. In Fig. 6, the tracking performance of the algorithms is investigated when the one group-sparse channel is used for first 150,000 iteration and the two group-sparse channel is considered for the following 150,000 iterations, and the other parameters remain unchanged. It is seen that the proposed RGS-PAP

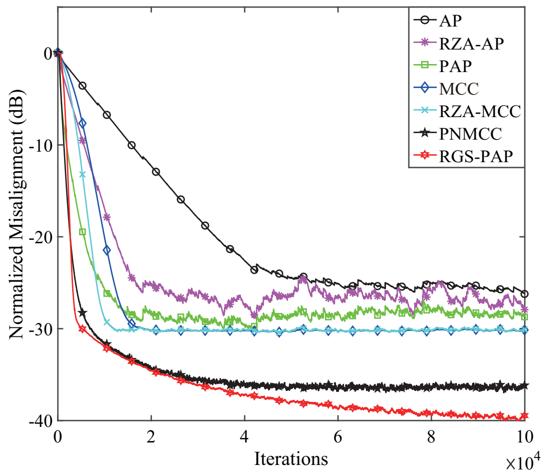

Fig. 2: Performance of different AF algorithms with one-group channel. Input signal: white Gaussian input signal.

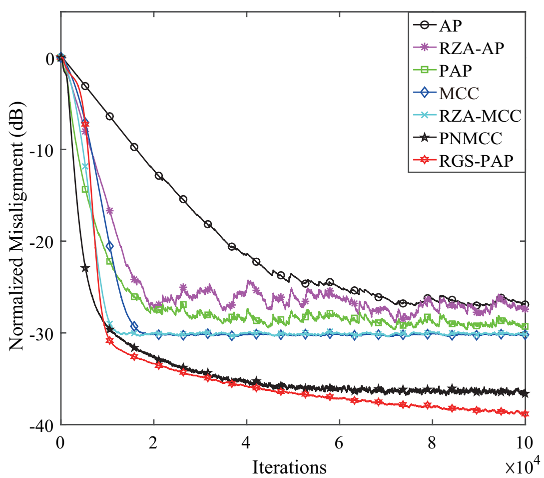

Fig. 3: Performance of different AF algorithms with two-group channel. Input signal: white Gaussian input signal.

algorithm provides the faster convergence and lowest NM.

C. Effects of SNR *on the RGS-PAP algorithm*

Fig.7 shows the steady-state NM performance against SNR for the colored input and the two-group channel. Herein, $\mu_{\mathrm{AP}} = 0.035$, $\mu_{\mathrm{RZA-AP}} = 0.055$, $\mu_{\mathrm{PAP}} = 0.01$, $\mu_{\mathrm{MCC}} = 0.0008$, $\mu_{\mathrm{RZA-MCC}} = 0.009$, $\mu_{\mathrm{PAP}} = 0.25$ and $\mu_{\mathrm{RGS-PAP}} = 0.01$ are selected and the other parameters are the same as in the previous experiment. Simulation results indicate that the performance of the proposed RGS-PAP algorithm is superior to those of the other algorithms.

V. Conclusion

In this paper, aiming at the problem that the MSE criterion is not suitable for impulsive noise environments, we use the MCC to modify the cost function of the AP algorithm, and then incorporate a mixed-norm into the modified cost function to exploit the group-sparse structure of the channel.

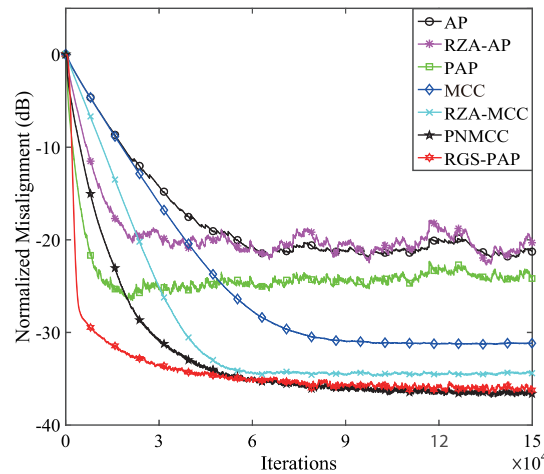

Fig. 4: Performance of different AF algorithms with one-group channel. Input signal: colored input signal.

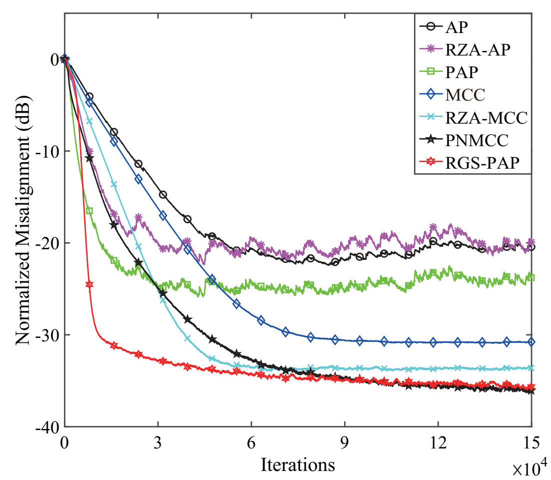

Fig. 5: Performance of different AF algorithms with two-group channel. Input signal: colored input signal.

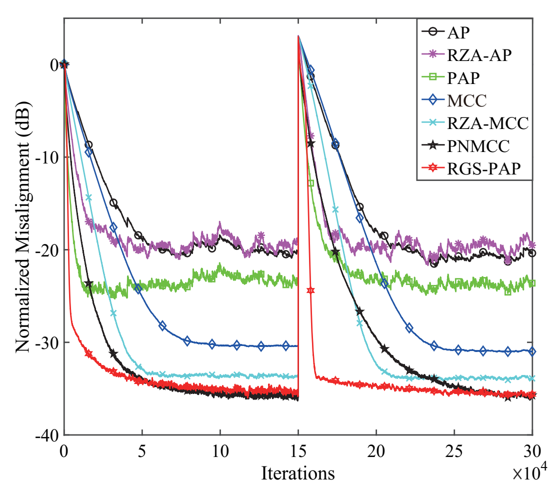

Fig. 6: Tracking performance of the algorithms.

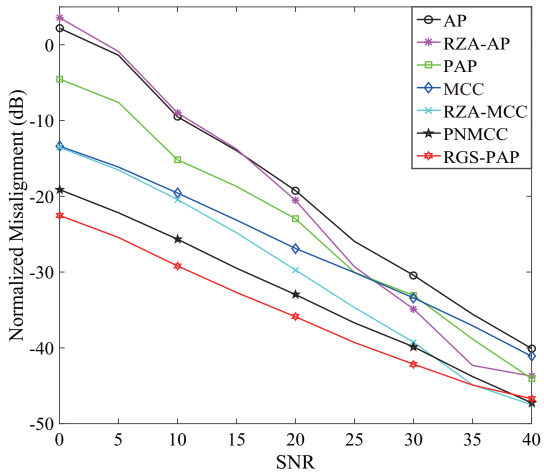

Fig. 7: Effect of SNR on the steady-state NM.

The developed RGS-PAP algorithm is investigated in various simulation experiments. The experimental results show that the proposed RGS-PAP algorithm has a faster convergence and smaller NM in $\alpha-$stable noise environments.

Acknowledgment

This work of Y. Li is funded by the Foundational Research Funds for the Central Universities (HEUCFG201829, 3072019CFG0801), Key Research and Development Program of Heilongjiang (GX17A016), the China Postdoctoral Science Foundation (2017M620918), and Z. Jiang is supported by the International Exchange Program of Harbin Engineering University for Innovation-oriented Talents Cultivation, and the work of Y. Zakharov is partly supported by the U.K. Engineering and Physical Sciences Research Council under Grant EP/P017975/1 and Grant EP/R003297/1.

References

[1] P. S. R. Diniz, Adaptive filtering: algorithms and practical implementation. 4th ed., Springer, New York, USA, 2013.

[2] L. Vuokko, V. -M. Kolmonen, J. Salo, and P. Vainikainen, "Measurement of large-scale cluster power characteristics for geometric channel models," IEEE Trans. on Antennas and Propagation, vol. 55, no. 11, pp. 3361-3365, 2007.

[3] 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.

[4] O. Oyerinde, and S. H. Mneney, "Review of channel estimation for wireless communication systems," IETE Technical Review, vol. 29, no. 4, pp. 282-298, 2012.

[5] P. C. Carrascosa , and M. Stojanovic, "Adaptive channel estimation and data detection for underwater acoustic MIMO-OFDM systems," IEEE Journal of Oceanic Engineering, vol. 35, no. 3, pp. 635-646, 2010.

[6] B. Widrow, and S. D. Stearns, Adaptive signal processing, Prentice Hall, New Jersey, USA, 1985.

[7] S. C. Douglas, and T. H. -Y. Meng, "Normalized data nonlinearities for LMS adaptation," IEEE Trans. on Signal Process., vol. 42, no. 6, pp. 1352-1365, June 1994.

[8] A. Gonzalez, M. Ferrer, F. Albu, and M. de Diego, "Affine projection algorithms: evolution to smart and fast algorithms and applications," 2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO), Bucharest, Romania, 2012.

[9] S. G. Sankaran, and A. A. L. Beex, "Convergence behavior of affine projection algorithms," IEEE Trans. on Signal Process., vol. 48, no. 4, pp. 1086-1096, Apr. 2000.

[10] A. Singh, and J. C. Principe, "Using Correntropy as a cost function in linear adaptive filters," Int. Joint Conf. on Neural Networks, pp. 2950-295, Atlanta, GA, USA, June 2009.

[11] Y. Li and M. Hamamura, "Smooth approximation l0-norm constrained affine projection algorithm and its applications in sparse channel estimation," The Scientific World Journal, vol.2014, Article ID: 937252, 2014.

[12] Y. Li and M. Hamamura, "Zero-attracting variable-step-size least mean square algorithms for adaptive sparse channel estimation," International Journal of Adaptive Control and Signal Processing, vol.29, no.9, pp.1189-1206, 2015.

[13] Y. Li, Y. Wang, R. Yang, F. Albu, "A soft parameter function penalized normalized maximum correntropy criterion algorithm for sparse system identification," Entropy, vol.19, no.1, ID:45, pp.1-16, 2017.

[14] B. Chen, J. Wang, H. Zhao, N. Zheng, and J. C. Principe, "Convergence of a fixed-point algorithm under maximum correntropy criterion," IEEE Signal Process. Letters, vol. 64, no. 13, pp. 3376-3387, July 2016.

[15] N. Richard, U. Mitra, "Sparse channel estimation for cooperative underwater communications: a structured multichannel approach," IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'08), Las Vegas, NV, USA, 2008.

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

[17] Y. Li, Y. Wang, and T. Jiang, "Sparse least mean mixed-norm adaptive filtering algorithms for sparse channel estimation applications," International Journal of Communication Systems, vol. 30, no. 8, pp. 1-14, Aug. 2016.

[18] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," IEEE Trans. on Speech and Audio Process., vol. 8, no. 5, pp. 508-518, Sept. 2000.

[19] H. Deng, and M. Doroslovački, "Improving convergence of the PNLMS algorithm for sparse impulse response identification," IEEE Signal Process. Letters, vol. 12, no. 3, pp. 181-184, Mar. 2005.

[20] Y. Chen, Y. Gu, and A. O. Hero, "Sparse LMS for system identification," IEEE Int. Conf. on Acoustics, Speech and Signal Processing, pp. 3125-3128, Taipei, Taiwan, Apr. 2009.

[21] 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, Apr. 2016.

[22] Y. Li, and M. Hamamura, "An improved proportionate normalized least-mean-square algorithm for broadband multipath channel estimation," The Scientific World Journal, vol. 2014, no. 8, pp. 1-9, Mar. 2014.

[23] T. Gänsler, S. L. Gay, M. M. Sondhi, and J. Benesty, "Double-talk robust fast converging algorithms for network echo cancellation," IEEE Trans. on Speech and Audio Process., vol. 8, no. 6, pp. 656-663, Nov. 2000.

[24] P. J. McLane, P. H. Wittke, P. K.-M. Ho, and C. Loo, "PSK and DPSK trellis codes for fast fading, shadowed mobile satellite communication channels," IEEE Trans. on Communications, vol. 36, no. 11, pp. 1242-1246, Nov. 1988.

[25] S. Jiang, and Y. Gu, "Block-sparsity-induced adaptive filter for multi-clustering system identification," IEEE Trans. on Signal Process., vol. 63, no. 20, pp. 5318-5330, Oct. 2015.

[26] J. Liu, and S. L. Grant, "Proportionate adaptive filtering for block-sparse system identification," IEEE/ACM Trans. on Audio, Speech, and Language Process., vol. 24, no. 4, pp. 623-630, Apr. 2016.

[27] 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/TCSII.2019.2891654.

[28] Y. Li, Z. Jiang, Z. Jin, X. Han, J. Yin, "Cluster-sparse proportionate NLMS algorithm with the hybrid norm constraint," IEEE Access, vol.6, pp.47794-47803, 2018.

[29] J. Liu, S. L. Grant, "Block sparse memory improved proportionate affine projection sign algorithm," Electronics Letters, vol. 51, no. 24, pp.2001-2003, Nov. 2015.

[30] J. Benesty, C. Paleologu, and S. Ciochina, "Proportionate adaptive filters from a basis pursuit perspective," IEEE Signal Process. Letters, vol. 17, no. 12, pp. 985-988, Dec. 2010.

[31] J. Benesty, C. Paleologu, and S. Ciochina, "On regularization in adaptive filtering," IEEE Trans. on Audio, Speech, Language Process., vol. 19, no. 6, pp. 1734-1742, Aug. 2011.