Convex Combination of Two Adaptive Filters with Normalized Median Wilcoxon Approach

Satyakam Baraha*1, Saurav Gupta*2 and Ajit Kumar Sahoo*3
*Department of Electronics & Communication Engineering, National Institute of Technology Rourkela, India
Email: satyakambaraha93@gmail.com1, greater.saurav@gmail.com2 and ajitsahoo@nitrkl.ac.in3

***Abstract*—Adaptive system identification is an active area of research in the field of signal processing due to the fact that it updates the unknown parameter of the filter using a suitable adaptive algorithm. Due to great advancements in digital signal processors in terms of high speed and low power consumption, adaptive algorithms are able to calculate the parameters of interest very fast with reduced complexity. Taking one of the practical aspect of adaptive algorithm such as convergence speed, many algorithms have been found in the literature that quickly converge to some steady state value. But presence of outliers in the data limit the performance of these algorithms. Though minimum Wilcoxon norm [1] has been proposed to overcome the effect of outliers, yet the speed of convergence can still be improved. A convex combination of two adaptive filters have been taken into account to tackle the above limitation. Results show that convex combination of filters work very efficiently against conventional adaptive algorithms, giving faster convergence. But in the presence of outliers its effectiveness is not up to the mark. A normalized median based approach is proposed, which is applied along with the convex combination in order to improve the performance. Simulation results defend the above statement and verify the authenticity of the research.**

***Index Terms*—LMS, Wilcoxon norm, block LMS, convex combination, convergence.**

I. Introduction

Wide use of least mean square (LMS) algorithm and its variants are basically due to its low complexity and convergence in stationary environment [2]. The conventional LMS algorithm is based on minimizing the quadratic norm of error. But when outliers come into picture, its robustness gets affected. In practical applications the data are not in an uniform range. It is highly possible that there are some observations which lies at an abnormal range present in the data. These values must be handled carefully otherwise they affect the overall system in terms of convergence. Majhi et al. [1] stated an algorithm based on minimum Wilcoxon norm giving robust and improved performance against such outliers. The merit of their work is that it can handle upto 40% outliers. But the limitation of the stated problem is the speed of convergence. Ban et al. [3] proposed normalized minimum Wilcoxon norm and affine projection algorithm (APA) Wilcoxon norm, which is an extension of the minimum Wilcoxon norm. Sahoo et al. [4] proposed two new schemes sign-regressor and sign-sign Wilcoxon and showed that both gives better results compared to conventional Wilcoxon algorithm in terms of speed of convergence. Dash et al. [5] carried out the analysis of outliers for system identification and showed that minimum Wilcoxon norm has better robustness against the outliers. Sananda et al. [6] proposed Wilcoxon norm for distributed adaptive networks. They carried out the steady state analysis using asymptotic linearity rank test for diffusion based distributed systems. Though Wilcoxon based adaptive algorithms give better performance against outliers, but its convergence speed can still be improved. Gracia et al. [7] proposed the performance analysis of convex combination of two adaptive filters and showed that the steady state performance is much better than the conventional approach. Hence, in this work we have employed a convex combination of two adaptive filters in order to enhance the convergence rate. We have taken two possible scenarios, i.e. system without outliers and with outliers. In the absence of outliers we have employed convex combination in LMS and its variants to verify that the speed of convergence is much faster than the conventional algorithms with single filter. For presence of outliers in the system we have taken minimum Wilcoxon norm and compared the result with its convex combination. Results show that in presence of outliers the convergence speed of minimum Wilcoxon norm and its convex combination are almost same. Hence, a normalized median based approach is proposed which is incorporated along with the convex combination resulting in very fast convergence in presence of 10 to 40 % outliers.

The orientation of this article is as follows. Section II gives a brief introduction of LMS algorithms and its variants. The concept of convex combination of adaptive filters & a short description of the minimum Wilcoxon norm also have been discussed in this section. The proposed technique combining the convex combination along with the normalized median approach has been given in section III. In section IV, an exhaustive comparison analysis of proposed technique and the conventional techniques have been given in the presence and absence of outliers and its MATLAB simulation results have been discussed. The concluding remarks of the proposed work is given in section V.

The notations used in this article are as follows: bold letters are used to specify the column vectors, bold italic symbols denote the matrix quantity and $\Re$ has been used for the set of real numbers. Here, **W** and **X** are the weight and input vectors, μ is the step size, d is the desired signal, y is the response of the system and e is the error. It is important to note that the input **X** is column vector in the absence of outliers, but

978-1-5386-9279-0/19/$31.00 ©2019 IEEE

in presence of outliers as it operates in a block, hence $\boldsymbol{X}$ is a matrix.

II. System Identification and Convex Combination Scheme

To identify an unknown system from input output signal is known as system identification. Both the unknown system and adaptive filter are driven by the same input. Adaptive filter adjusts its weights so that it can match the weights of the unknown plant. Upon convergence, if the relationship of input and output for both the system and the adaptive model are same (i.e. the error is minimal), then it can be said that the adaptive system is a model of the unknown system.

A. Adaptive Algorithms

Different algorithms which are used to verify the authenticity of the result, and are briefly over-viewed in this subsection. We have mainly focused on the stochastic gradient approach, i.e. LMS algorithm and its different variants. The different adaptive algorithms are

1) **LMS Algorithm :** LMS algorithm [8] is widely used because of its low computational complexity and it doesn't need the statistical information such as R_{xx} and R_{xy} to update the weights. It is given by

$$\mathbf{W}(n) = \mathbf{W}(n-1) + \mu\mathbf{X}(n)e(n). \quad (1)$$

2) **NLMS Algorithm :** The limitation of the LMS algorithm is that if the inputs are scaled, then its performance decreases [9]. Hence NLMS [8] is given by

$$\mathbf{W}(n) = \mathbf{W}(n-1) + \mu[\frac{\mathbf{X}(n)(e(n))}{\epsilon + \|\mathbf{X}(n)\|_2^2}]. \quad (2)$$

3) **Sign-error LMS Algorithm :** The sign-error LMS [4] is given by

$$\mathbf{W}(n) = \mathbf{W}_(n-1) + \mu\mathbf{X}(n)sgn[e(n)]. \quad (3)$$

4) **Leaky LMS Algorithm :** Gupta et al. and Modalavalasa et al. [10], [11] stated the leaky LMS algorithm which is given by

$$\mathbf{W}(n) = (1-\gamma\mu)\mathbf{W}_(n-1) + 2\mu e(n)\mathbf{X}(n). \quad (4)$$

B. Convex Combination of Two Adaptive Filters

The adaptive convex combination scheme has been shown in Fig. 1, from which the overall output of the filter [7] can be given as

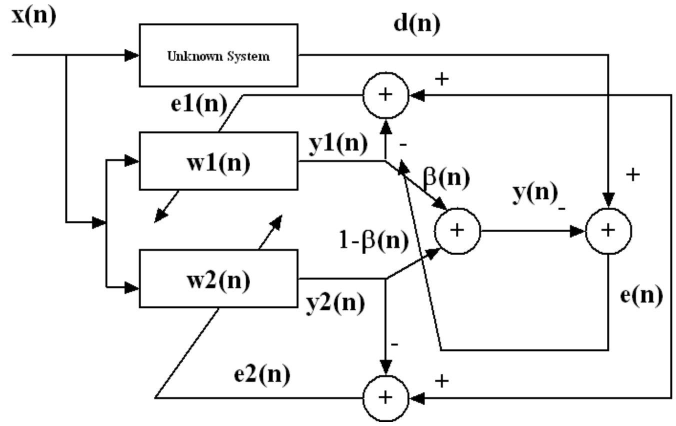

Fig. 1: System identification using adaptive convex combination of two transversal filters

$$y(n) = \beta(n)y_1(n) + [1-\beta(n)]y_2(n), \quad (5)$$

where $y_i(n)$; $\forall i = 1,2$ are the outputs of the two transversal filters at instant n, i.e. $y_i(n) = w_i^T(n)x(n), \forall i = 1,2$. $w_i^T(n)$ are the weights of individual filters and $x(n)$ is the input vector. $\beta(n)$ is a scalar parameter such that $0 \leq \beta \leq 1$. $w_i(n), \forall i = 1,2$ will give best results if $\beta(n)$ is tuned properly in each iteration.

The individual filter weights are updated independently with the help of general transversal scheme as follows

$$\mathbf{W}_i(n+1) = f_i(\mathbf{W}_i(n), d(n), \mathbf{X}(n), e(n)), \quad (6)$$

where the desired signal is given by $d(n)$, and $f_i(.)$ is any suitable adaptive algorithm. The overall weights get updated using the following equation

$$\mathbf{W}(n) = \beta(n)\mathbf{W}_1(n) + [1-\beta(n)]\mathbf{W}_2(n). \quad (7)$$

The parameter $\beta(n)$ is given by the following sigmoid function

$$\beta(n) = sigm[c(n)] = \frac{1}{1+e^{-c(n)}}, \quad (8)$$

which is used to combine both the outputs and $c(n)$ is a variable. However $\beta(n)$ gets modified indirectly by adapting $c(n)$ using a gradient-descent method to minimize the overall error [12]

$$\begin{aligned} c(n+1) =& c(n) - \frac{\mu}{2}\frac{\partial e^2(n)}{\partial c(n)} \\ =& c(n) - \frac{\mu}{2}\frac{\partial e^2(n)}{\partial e(n)}\frac{\partial e(n)}{\partial y(n)}\frac{\partial y(n)}{\partial \beta(n)}\frac{\partial \beta(n)}{\partial c(n)} \\ =& c(n) + \mu e(n)[y_1(n) - y_2(n)]\beta(n)[1-\beta(n)]. \end{aligned} \quad (9)$$

The adaptation in (5) will stop when $\beta(n)$ value is either 0 or 1. To overcome this problem the value of $c(n)$ restricted in the range $[-c^+, c^+]$, which in turn limits the range of $\beta(n)$ in the range $[1-\beta^+, \beta^+]$.

Depending on the value of $\beta(n)$, the response of the filter will vary. Hence the modified response of the filter can be given by

$$y_u(n) = \beta_u(n)y_1(n) + [1-\beta_u(n)]y_2(n), \quad (10)$$

where

$$\beta_u(n) = \begin{cases} 1; & c(n) \geq c^+ - \epsilon \\ \beta(n); & -c^+ + \epsilon < c(n) < c^+ - \epsilon \\ 0; & c(n) \leq -c^+ + \epsilon \end{cases}, \quad (11)$$

with ϵ being a small positive quantity.

C. Minimum Wilcoxon Norm

Unlike conventional algorithms, minimum Wilcoxon norm operates in block. Hence the desired signal and error signal in this case are column vectors. Let the output of the adaptive filter at pth input and qth experiment is given by

$$\mathbf{y}(p,q) = \mathbf{W}(q)^T \boldsymbol{X}_p. \quad (12)$$

The error is given by

$$\mathbf{e}(p,q) = \mathbf{d}(p,q) - \mathbf{y}(p,q). \quad (13)$$

Minimum Wilcoxon norm of error is given by

$$\psi(q) = \sum_{p=1}^{K} \mathbf{s}(p,q)\mathbf{e}_p(q), \quad (14)$$

where $s(p,q)$ is the score function such that $\phi(k) : [0,1] \rightarrow R$, satisfying

$$\int_0^1 \phi^2(k)dk < \infty, \quad (15)$$

where

$$\phi(k) = \sqrt{12}(k - 0.5). \quad (16)$$

Taking $k = \frac{R\{e(p,q)\}}{K+1}$, the score $\phi(k)$ is given by

$$\mathbf{s}(p,q) = \sqrt{12}(\frac{R\{\mathbf{e}(p,q)\}}{K+1} - 0.5); 1 \leq p \leq K, \quad (17)$$

where $R\{\mathbf{e}(p,q)\}$ indicates the rank of error such that

$$\mathbf{e}_p(q) \leq \mathbf{e}_p(q+1); 1 \leq k \leq K. \quad (18)$$

Finally the weight update equation is given by

$$\begin{aligned} \mathbf{W}(q+1) =& \mathbf{W}(q) - \mu \frac{\partial \psi(q)}{\partial \mathbf{W}} \\ =& \mathbf{W}(q) - \mu \frac{\partial \psi(q)}{\partial \mathbf{e}_p(q)} \frac{\partial \mathbf{e}_p(q)}{\partial \mathbf{y}(p,q)} \frac{\partial \mathbf{y}(p,q)}{\partial \mathbf{W}} \\ =& \mathbf{W}(q) - \mu \sum_{p=1}^{K} \boldsymbol{X}(p)\mathbf{s}(p,q). \end{aligned} \quad (19)$$

III. Proposed Method

Ban et al. [3] proposed Normalized Wilcoxon given by

$$\mathbf{W}_i = \mathbf{W}_{i-1} + \mu \boldsymbol{X}_i(\boldsymbol{X}_i\boldsymbol{X}_i^{'})^{-1}\mathbf{s}_i. \quad (20)$$

From (14), (16) and (17) the minimum Wilcoxon norm can be expressed by the following form [4]

$$|\psi| = \sum_{i=1}^{K} \phi(e_i)e_i. \quad (21)$$

From (21) replace e_i by $(e_i - med(e_i))$, then it will become

$$\begin{aligned} |\psi| &= \sum_{i=1}^{K} \phi(e_i)(e_i - med(e_i)) \\ &= \sum_{i=1}^{K} [\frac{\phi(e_i)}{(e_i - med(e_i))}](e_i - med(e_i))^2 \\ &= \sum_{i=1}^{K} (\omega_i)(e_i - med(e_i))^2 \\ &= \sum_{i=1}^{K} (\omega_i)((d_i - \boldsymbol{X}_i^T\mathbf{W}) - med(e_i))^2 \\ &= \sum_{i=1}^{K} (\omega_i)((d_i - med(e_i)) - \boldsymbol{X}_i^T\mathbf{W})^2 \\ &= \sum_{i=1}^{K} (\omega_i)(\bar{d}_i - \boldsymbol{X}_i^T\mathbf{W})^2 \\ &= \sum_{i=1}^{K} (\sqrt{\omega_i}\bar{d}_i - \sqrt{\omega_i}\boldsymbol{X}_i^T\mathbf{W})^2 \\ &= \sum_{i=1}^{K} (\bar{d}_i^{\,w} - (\boldsymbol{X}_i^w)^T\mathbf{W})^2 \end{aligned}, \quad (22)$$

where $\boldsymbol{X}_i^w = \sqrt{\omega_i}\boldsymbol{X}_i$ and $\bar{d}_i^{\,w} = \sqrt{\omega_i}\bar{d}_i$. Eq. (22) indicates the new formulation of minimum Wilcoxon norm. Incorporating the changes, $\boldsymbol{X}$ in (20) we will get the final update equation which is given by

$$\mathbf{W}_i = \mathbf{W}_{i-1} + \mu \boldsymbol{X}_i^w(\boldsymbol{X}_i^w\boldsymbol{X}_i^{w'}))^{-1}\mathbf{s}_{conv}. \quad (23)$$

It would be interesting to see the effect of convex combination of two filters with minimum Wilcoxon norm in the presence of outliers in the data. The above method is employed along with convex combination scheme in order to increase the speed of convergence. Eq. (23) represents the weight update equation of the proposed method. In Eq. (23) $\mathbf{s}_{conv}$ indicates the score value of the convex combination and is calculated using individual $\mathbf{s}_i$, i.e.

$$\mathbf{s}_{conv} = \beta(n)\mathbf{s}_1 + [1 - \beta(n)]\mathbf{s}_2, \quad (24)$$

where $\mathbf{s}_1$ and $\mathbf{s}_2$ are the individual scores of the adaptive filters. The system is modeled as

$$d(n) = \mathbf{W}^T(n)\mathbf{X}(n) + \eta(n) + o(n); n = 1,2,...,N. \quad (25)$$

$\mathbf{W}\epsilon\Re^L$, $\mathbf{X}\epsilon\Re^L$ are the weight and input column vectors respectively. The order of the system is given by L. η indicates the additive white Gaussian noise (AWGN) while o represents the outliers present in the desired data. d is the desired response of the unknown system.

The output of individual transversal filters are given by

$$y_i(n) = \mathbf{W}_i^T(n)\mathbf{X}(n); i = 1,2. \quad (26)$$

The respective errors are given by

$$e_i(n) = d(n) - y_i(n); i = 1,2. \quad (27)$$

and the corresponding score value $\mathbf{s}_i(n)$ is calculated using (17). The Wilcoxon norm $\psi_i(n)$ for the two score values are calculated using (14). The individual weights of the two adaptive filters are updated using (19). From (11) the value of $\beta_u(n)$ is calculated with respect to $c(n)$. The overall output of the filter $y(n)$ is calculated using (5). The final error is calculated as

$$e(n) = d(n) - y(n); i = 1, 2. \quad (28)$$

The overall weight $\mathbf{W}$ is updated using (7). For the proposed method, the updated weight is modified using normalized median approach given in (22) and hence the new weight update equation is given by Eq. (23). Weight update using proposed method is helpful for getting faster convergence. Now $c(n)$ is updated using (9) and the value of $\beta(n)$ is updated using (8).

IV. Results and Analysis

This section contains the results and analysis of the proposed method and the discussions made above. The simulations are carried out using MATLAB software. Results shown in this section depicts two possible scenarios. First, in the absence of outliers the LMS algorithm and its variants are employed with convex combination. Second, in the presence of up to 40 % outliers, convex combination is applied to minimum Wilcoxon norm. The weight of the unknown system is taken as

$$\mathbf{W}_0 = [0.26\ 0.93\ 0.26]^T \quad (29)$$

A. Absence of outliers

In the absence of outliers 20000 number of input data is generated within the range of random value $(-0.5, 0.5)$. The response of the unknown system is calculated using the tapped delay structure as given in (25). The AWGN noise is given by η. The output have been obtained by conducting the average of total 100 number of Monte Carlo runs. Fig. 2 (a) represents the MSE plot of the linear system given in Eq. (25) using the adaptive algorithms given in Eq. (1), (2), (3) and (4) respectively. We can clearly see that NLMS is converging faster compared to LMS and leaky LMS. The converging speed of sign error LMS is fastest among all the four.

In Fig. 2 (b) the MSE of respective convex combinations of the above algorithms have been plotted. The same pattern as we got in Fig. 2 (a) can be observed in Fig. 2 (b), but with fast convergence. Convex sign error LMS is fastest, while convex LMS is the slowest. Convex NLMS is faster compared to convex LMS and convex leaky LMS.

The MSE plot of convex combination scheme of LMS and NLMS have been compared with their conventional algorithms have been plotted in Fig. 2 (c). The convex scheme is faster in achieving the steady-state compared to the existing methods, so is the case with sign-error and leaky LMS.

B. Presence of outliers

The outliers o is present along with the AWGN η, which makes it difficult for the system to converge. Hence the data is processed in block. The value of outliers is random value fixed between $(-V, V)$ where V is an observation lying far away from the existing values present in the data. Percentage of outliers is decided by calculating $\frac{\text{Number of outlier instances}}{\text{Block size}} \times 100$. Simulations have been carried out by taking the values of outliers in the range $10 \leq |o| \leq 20$, which is far away from the values present in the data lying in the range (-0.5, 0.5). In simulations normalized mean square deviations (NMSD) in dB is plotted with respect to number of iterations. The NMSD is calculate using $10log_{10}\frac{\|\mathbf{W}_0-\mathbf{W}\|_2^2}{\|\mathbf{W}_0\|_2^2}$. Fig. 3 (a), (b) and (c) shows the simulations for three algorithms, i.e. minimum Wilcoxon norm, convex minimum Wilcoxon norm and the proposed convex normalized median Wilcoxon norm for 20%, 30% and 40% outliers respectively. It reflects that the effect of convex combination in presence of outliers is same as the conventional minimum Wilcoxon norm. But after incorporating the proposed changes as given in (22) and (23), the convergence speed is increased drastically, increasing the accuracy of the result. Though convergence speed is getting improved by the proposed method, the quality of the aforementioned method can be proved by calculating the % deviation and by comparing our result with rest of the algorithms, in presence of outliers. Table I shows the value of estimated weights for the initial weight given in (29) using block LMS, Wilcoxon norm, convex Wilcoxon norm and the proposed convex normalized median Wilcoxon norm. It clearly depicts that in presence of outliers the block LMS never converges. The estimated values are far away from the original value and the % deviation is very high. Wilcoxon norm gives standard results and work fine against outliers. The convex combination of minimum Wilcoxon norm is also not improving the results up to standard. But when the convex combination is employed in normalized minimum Wilcoxon norm, it gives efficient results. The estimated values are very close to the original value and the % deviation is also very less.

V. Conclusion

This research has been focused to study the effect of convex combination on adaptive algorithms both in presence and absence of outliers. In the absence of outliers it increases the speed of convergence of existing algorithms. In the presence of outliers it doesn't improve the converging speed up to the mark. By incorporating the changes and using normalized median method along with convex combination, the convergence speed of the system increases significantly, nullifying the effect of outliers. Another interesting result is when the amount of outliers increase, the deviation still remains less. Stability and steady state analysis of the proposed method can be considered as the future work of this research.

References

[1] B. Majhi, G. Panda, and B. Mulgrew, "Robust identification using new wilcoxon least mean square algorithm," *Electronics letters*, vol. 45, no. 6, pp. 334–335, 2009.

[2] R. K. Thenua and S. Agarwal, "Simulation and performance analysis of adaptive filter in noise cancellation," *International Journal of Engineering Science and Technology*, vol. 2, no. 9, pp. 4373–4378, 2010.

[3] S. Ban and S. Kim, "Wilcoxon adaptive algorithms for robust identification," *Electronics letters*, vol. 45, no. 18, pp. 958–959, 2009.

[4] U. K. Sahoo, G. Panda, and B. Mulgrew, "Sign-regressor wilcoxon and sign-sign wilcoxon," in *Advances in Recent Technologies in Communication and Computing (ARTCom), 2010 International Conference on*, pp. 35–39, IEEE, 2010.

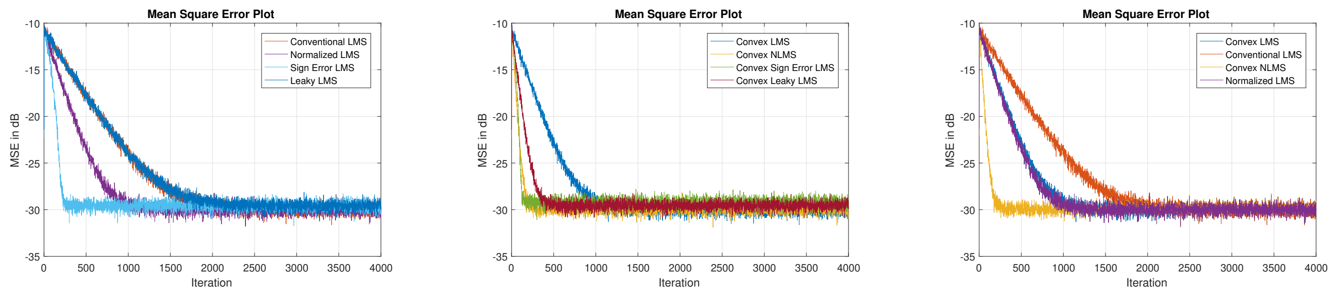

Fig. 2: Comparison of MSE of (a) LMS algorithm & its variants (b) convex combination employed on the respective algorithms (c) convex combination with existing algorithms

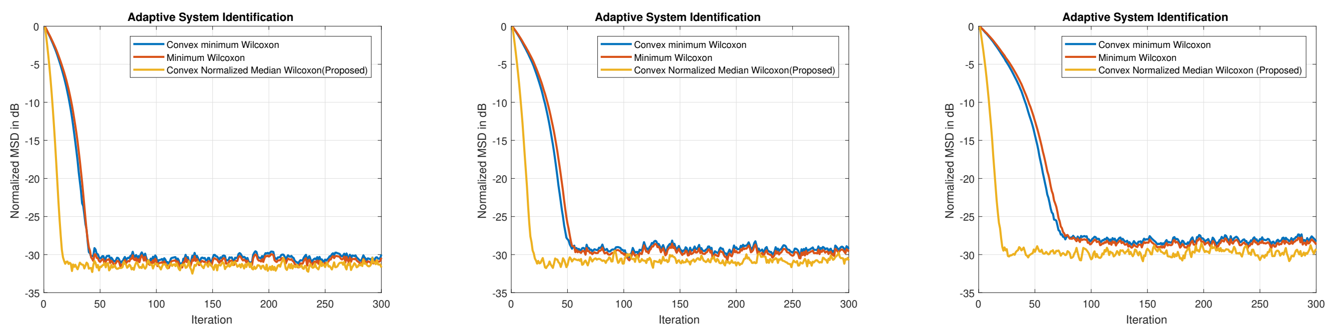

Fig. 3: Comparison of Normalized MSD of conventional Wilcoxon norm, convex Wilcoxon norm & convex normalized median Wilcoxon with (a) 20% outliers (b) 30% outliers (c) 40% outliers

TABLE I: Comparison of weight parameter for different algorithms in presence of outliers at 30 dB SNR

Outliers	Parameter ($\mathbf{W}_0$)	Block LMS Weight	Block LMS %D	Wilcoxon Norm Weight	Wilcoxon Norm % D	Convex Wilcoxon Weight	Convex Wilcoxon %D	Convvex Normalized Median Wilcoxon(Proposed) Weight	Convvex Normalized Median Wilcoxon(Proposed) %D
20%	0.26	0.6231	58.27	0.2559	1.6	0.2400	8.33	**0.2608**	**0.30**
	0.93	0.8101	14.8	0.9481	1.9	0.9541	2.52	**0.9218**	**0.88**
	0.26	0.1443	80.18	0.2646	1.73	0.2508	3.66	**0.2579**	**0.81**
40%	0.26	1.7508	85.14	0.2875	9.56	0.2531	2.72	**0.2631**	**1.17**
	0.93	-0.0615	1.6122e+03	0.9537	2.48	0.9172	1.39	**0.9358**	**0.61**
	0.26	1.7887	85.51	0.2795	6.97	0.2577	0.89	**0.2613**	**0.49**

[5] S. Dash and M. N. Mohanty, "Analysis of outliers in system identification using wlms algorithm," in *Computing, Electronics and Electrical Technologies (ICCEET), 2012 International Conference on*, pp. 802–806, IEEE, 2012.

[6] S. Kumar, U. K. Sahoo, A. K. Sahoo, and D. P. Acharya, "Diffusion minimum-wilcoxon-norm over distributed adaptive networks: Formulation and performance analysis," *Digital Signal Processing*, vol. 51, pp. 156–169, 2016.

[7] J. Arenas-García, A. R. Figueiras-Vidal, and A. H. Sayed, "Mean-square performance of a convex combination of two adaptive filters," *IEEE transactions on signal processing*, vol. 54, no. 3, pp. 1078–1090, 2006.

[8] S. Haykin and B. Widrow, *Least-mean-square adaptive filters*, vol. 31. John Wiley & Sons, 2003.

[9] J. Dhiman, S. Ahmad, and K. Gulia, "Comparison between adaptive filter algorithms (lms, nlms and rls)," *International Journal of Science, Engineering and Technology Research (IJSETR)*, vol. 2, no. 5, pp. 1100–1103, 2013.

[10] S. Gupta, A. K. Sahoo, and U. K. Sahoo, "Parameter estimation of wiener nonlinear model using least mean square (lms) algorithm," in *Region 10 Conference, TENCON 2017-2017 IEEE*, pp. 1399–1403, IEEE, 2017.

[11] S. Modalavalasa, U. K. Sahoo, and A. K. Sahoo, "Incremental modified leaky lms," 2017.

[12] Y. Zhang and J. A. Chambers, "Convex combination of adaptive filters for a variable tap-length lms algorithm," *IEEE Signal Processing Letters*, vol. 13, no. 10, pp. 628–631, 2006.