Speech Enhancement by Modified Convex Combination of Fractional Adaptive Filtering

S. Ghalamiosgouei* and M. Geravanchizadeh*[(C.A.)]

Abstract: This paper presents new adaptive filtering techniques used in speech enhancement system. Adaptive filtering schemes are subjected to different trade-offs regarding their steady-state misadjustment, speed of convergence, and tracking performance. Fractional Least-Mean-Square (FLMS) is a new adaptive algorithm which has better performance than the conventional LMS algorithm. Normalization of LMS leads to better performance of adaptive filter. Furthermore, convex combination of two adaptive filters improves its performance. In this paper, new convex combinational adaptive filtering methods in the framework of speech enhancement system are proposed. The proposed methods utilize the idea of normalization and fractional derivative, both in the design of different convex mixing strategies and their related component filters. To assess our proposed methods, simulation results of different LMS-based algorithms based on their convergence behavior (i.e., MSE plots) and different objective and subjective criteria are compared. The objective and subjective evaluations include examining the results of SNR improvement, PESQ test, and listening tests for dual-channel speech enhancement. The powerful aspects of proposed methods are their low complexity, as expected with all LMS-based methods, along with a high convergence rate.

Keywords: Adaptive Filters, Convex Combination of Adaptive Filters, Fractional Least-Mean-Squares, Least-Mean-Squares, Normalized Fractional Least-mean-squares, Speech Enhancement.

1 Introduction

Speech communication devices are often used in environments with high levels of ambient noise such as cars and public places. The noise picked up by microphones of the device can significantly impair the quality of the transmitted speech signal. When the intelligibility of the transmitted speech is also impaired, the device cannot be used in the desired way. It is therefore sensible to include a noise reduction pre-processor in such devices.

Numerous schemes have been proposed and implemented that perform speech enhancement under various constraints and/or assumptions and deal with different issues and applications [1].

Nowadays, adaptive algorithms represent one of the most frequently used computational tools for the processing of digital speech signals. As special case, dual-channel speech enhancement is one of the digital signal processing subjects which uses adaptive filtering. Such systems incorporate two microphones, in which one of the microphones receives noisy speech signal and the other one takes noise signal [2].

There are many types of adaptive filters which employ different schemes to adjust filter weights. Among all adaptive algorithms, Widrow and Hoff's Least-Mean-Squares (LMS) [3, 4] has probably become the most popular algorithm for its robustness, good tracking capabilities, and simplicity, both in terms of computational load and easiness of implementation. The main drawback of the "pure" LMS algorithm is that it is sensitive to the scaling of its input. To solve this problem, filter weights are normalized with the power of the input. This variant of the LMS algorithm is called Normalized Least-Mean-Squares (NLMS) [5, 6].

The concept of fractional order operators has been investigated extensively in recent years in various signal processing theories and techniques [7-10]. Recently, a new method based on the modification of LMS-based adaptive filters has been proposed, which uses the fractional order derivative of Mean-Square Error (MSE) together with the first order derivative
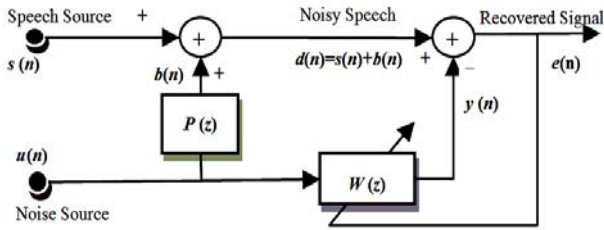
Fig. 1 Dual-channel speech enhancement structure.

[11]. In [12], a method based on Fractional Least-Mean-Square (FLMS) algorithm is presented to work with nonlinear time series prediction. More recent applications of FLMS in signal processing methods include echo cancellation problem [13], and parameter estimation of Input Nonlinear Control Autoregressive (INCAR) models [14].

The design of many adaptive filters requires a trade-off between convergence speed and steady-state mean-squares error. A faster (or slower) convergence speed yields a larger (or smaller) steady-state Mean-Square Deviation (MSD) and MSE. With this aspect, combinations of adaptive filters have recently attracted attention due to their ability to improve transient and steady-state performance of adaptive filters in stationary and non-stationary environments. So far, many structures of combination filters [15-16] have been proposed. One of interesting combinations of adaptive filters is the convex combination of two adaptive filters, also called component filters [16-18]. In convex combination, the output signals and the output errors of both filters are combined in such a way that the advantages of both component filters, namely, the rapid convergence of the fast filter and the reduced steady-state error from the slow filter, are retained. Recently, the convex combination of filtered-x algorithm has been employed in active noise control [19].

In order to improve further the performance of convex combinational filter, normalized convex combination of adaptive filters has been introduced. It is shown that the new update rule preserves the good features of the existing scheme and is more robust to changes in the filtering scenario [20].

In this paper, new convex combinational adaptive filtering techniques are proposed, in which normalization and fractional order features are employed, both in structures of component filters and in mixing strategy of the combinational scheme.

This paper is organized as follows. Section 2 describes the dual-channel speech enhancement system together with the techniques of LMS, NLMS, FLMS, and the structures of convex combination and normalized convex combination of adaptive filters. In Section 3, our proposed convex combinational adaptive methods are introduced. The different convex combinational schemes discussed include Convex

Combination of Normalized Fractional Least-Mean-Squares (CC-NFLMS), fractional convex combination and fractional normalized convex combination of component filters. Section 4 presents the experimental results and comparisons with traditional LMS-based adaptive filtering methods used in the context of speech enhancement. Concluding remarks are given in Section 5.

2 Background
2.1 Dual-channel Speech Enhancement
Fig. 1 shows the block diagram for a general two-channel enhancement system. The clean speech signal $s(n)$ is assumed to be present in only one channel, which is then corrupted by the background noise $b(n)$ to generate the noisy speech signal $d(n)$. The second channel has the reference noise signal $u(n)$ as input. The acoustic path transfer function between two sensors is given by $P(z)$. The adaptive filter $W(z)$ tries to estimate the acoustic path transfer function $P(z)$. As a result, the filter output $y(n)$ becomes an estimate of only noise present in $d(n)$. The output of the adaptive filter is given by

$$y(n) = \mathbf{w}(n)^H \mathbf{u}(n) \tag{1}$$

where $\mathbf{w}$ is the weight vector with length L.

The output of the structure, $e(n)$, will be an estimate of the clean speech signal $s(n)$. In order to obtain the optimal adaptive filter coefficients, $\mathbf{w}$, the following cost function is minimized:

$$J(n) = E\left[e(n)e^*(n)\right] = E\left[\left|e(n)^2\right|\right] \tag{2}$$

where E denotes the expectation operator.

2.2 LMS Algorithm
The LMS algorithm [4] makes the simplifying assumption that the expected value of the squared error is approximated by the squared error itself, i.e., $E\left\{\left|e(n)\right|^2\right\} \cong \left|e(n)\right|^2$. In vector notation, the LMS update relation becomes:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu e(n)\mathbf{u}(n) \tag{3}$$

where μ is the step size.

2.3 NLMS Algorithm
In the LMS algorithm, the adjustment applied to the tap-weight vector is directly proportional to the input vector, $\mathbf{u}(n)$. Therefore, when $\mathbf{u}(n)$ is large, the LMS filters suffer from a gradient noise amplification problem. To overcome this difficulty, the NLMS filter can be used [5, 6].

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\tilde{\mu}}{\delta + \|\mathbf{u}(n)\|^2}\mathbf{u}(n)e(n) \tag{4}$$

where $\|\mathbf{u}(n)\|^2$ is the power of input vector and $\delta > 0$.

2.4 FLMS Algorithm

In deriving the FLMS algorithm, fractional derivatives in addition to the first derivative should be used. The update relation for the *k-th* element of the weight vector in FLMS is given by [11]:

$$w_k(n+1) = w_k(n) - \mu_1 \frac{\partial J(n)}{\partial w_k} - \mu_f \frac{\partial^v J(n)}{\partial w_k^v} \quad (5)$$

where v ($0 < v < 1$) is a real number, μ_1 is the first-order step size, and μ_f is the fractional step-size.

Applying fractional derivative of order α [21] to the mean-square error (cost function (2)), gives:

$$\frac{\partial^v J(n)}{\partial w_k^v} = -2e(n)u(n-k)\frac{1}{\Gamma(2-v)}w_k^{1-v}(n) \quad (6)$$

where $\Gamma(.)$ denotes the gamma function. The final update relation for the weight vectors of the FLMS algorithm can be written as:

$$w_k(n+1) = \begin{cases} w_k(n) + (\mu_1 + \mu_f \dfrac{|w_k(n)|^{1-v}}{\Gamma(2-v)})e(n)u(n-k)w_k \geq 0 \\[2ex] w_k(n) + (\mu_1 - \mu_f \dfrac{|w_k(n)|^{1-v}}{\Gamma(2-v)})e(n)u(n-k), w_k < 0 \end{cases} \quad (7)$$

It is also noteworthy that from the standpoint of implementation, here, a modified version of the update rule is used as compared with that given in [11]. The Eq. (7) can be rewritten as follows:

$$w_k(n+1) = w_k(n) +$$
$$\left(\mu_1 + \mu_f \frac{|w_k(n)|^{1-v}}{\Gamma(2-v)} \text{sgn}(w_k(n)) \right) e(n)u(n-k) \quad (8)$$

where sgn(.) denotes the sign function.

2.5 Convex Combination of Two Adaptive Filters

The structure of convex combination of two adaptive filters is shown in Fig. 2 [17, 18]. The output of the parallel filter is:

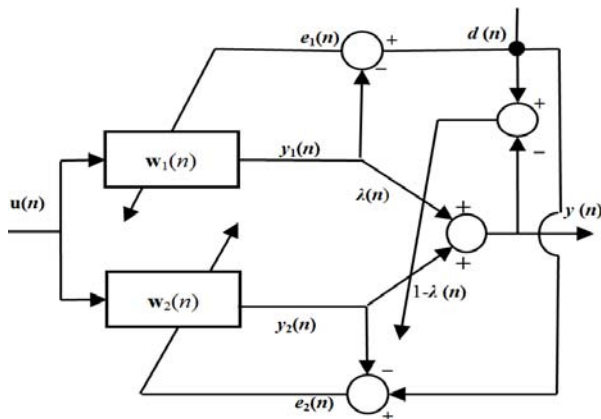$$y(n) = \lambda(n)y_1(n) + [1-\lambda(n)]y_2(n) \quad (9)$$

Fig. 2 Convex combination of two adaptive filters.

Table 1 The summary of convex combination algorithm [16].

1-	Initialization: $\mathbf{w}_1(0) = \mathbf{w}_2(0) = 0$; $a(0) = 0$; $\lambda(0)=0.5$; $\mu_1(0)$, $\mu_2(0)$, μ_a, $\mu_{\max}$, β, r
2-	Loop $n = 0,1,2,\ldots$ $\quad y_i(n) = \mathbf{w}_i^T(n)\mathbf{u}(n), \quad i=1,2$ $\quad e_i(n) = d(n) - y_i(n), \quad i=1,2$ $\quad \mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \mu_i(n)e_i(n)\mathbf{u}(n), \quad i=1,2$ $\quad y(n) = \lambda(n)y_1(n) + [1-\lambda(n)]y_2(n)$ $\quad e(n) = d(n) - y(n)$ $\quad a(n+1) = a(n) + \mu_a e(n)e(n)[y_1(n)-y_2(n)]\lambda(n)[1-\lambda(n)]$ $\quad \lambda(n+1) = \text{sgm}(a(n+1))$
2.1 -	if $\lambda(n+1) < 1-\beta$ $\quad \mu_i(n+1) = \mu_i(n)/r, i=1,2$ $\quad \mathbf{w}_1(n+1) = \mathbf{w}_2(n+1)$ $\quad a(n+1) = 0; \lambda(n+1) = 0.5$ Endif
2.2 -	if $(\lambda(n+1) > \beta$ and $r\mu_1(n) < \mu_{\max})$ $\quad \mu_i(n+1) = r\mu_i(n), i=1,2$ $\quad \mathbf{w}_2(n+1) = \mathbf{w}_1(n+1)$ $\quad a(n+1) = 0; \lambda(n+1) = 0.5$ Endif
2.3 -	if $(\lambda(n+1) > \beta$ and $r\mu_1(n) > \mu_{\max})$ $\quad \mu(n+1) = a^+; \lambda(n+1) = \beta$ Endif
	End

Here, $y_1(n) = \mathbf{w}_1^T \mathbf{u}_1(n)$ and $y_2(n) = \mathbf{w}_2^T \mathbf{u}_2(n)$ are the output of two parallel transversal filters at time n and $\lambda(n)$ is the mixing parameter limited in [0, 1].

The mixing parameter $\lambda(n)$ is updated via an auxiliary variable $a(n)$, which is defined as:

$$\lambda(n) = \text{sgm}[a(n)] \quad (10)$$

where sgm(.) is the sigmoidal function, defined as:

$$\text{sgm}[a(n)] = \frac{1}{1+e^{-a(n)}} \quad (11)$$

It is shown in [17] that if $\lambda(n)$ is chosen properly at each iteration, then the above combination extracts the best specifications of the individual filters, $\mathbf{w}_1(n)$ and $\mathbf{w}_2(n)$. The update equation for $a(n)$ is given by:

$$a(n+1) = a(n) - \frac{\mu_a}{2}\frac{\partial e^2(n)}{\partial a(n)}$$
$$= a(n) + \mu_a e(n)[y_1(n) - y_2(n)]\lambda(n)(1-\lambda(n)) \quad (12)$$

Table 1 presents the pseudo code of the convex combination algorithm.

2.6 Normalized Convex Combination of Adaptive Filters

The overall combinational scheme can be considered as a two-layer adaptive filter [15]. In the first layer, the two component filters operate independently of each other according to their own rules, while the second layer

consists of a filter with the input signal $e_2(n) - e_1(n)$ that minimizes the overall error.

The convex combinational filter proposed by [16] updates $a(n)$ by the standard LMS algorithm with the input $e_2(n) - e_1(n)$ and step-size $\mu_a \lambda(n)(1 - \lambda(n))$.

Considering the drawbacks of the conventional LMS algorithm discussed above, the parameter $a(n)$ can be updated efficiently with normalized LMS [20]:

$$a(n+1) = a(n) +$$

$$\frac{\mu_a}{p(n)} \lambda(n)[1 - \lambda(n)]e(n)[e_2(n) - e_1(n)] \qquad (13)$$

where

$$p(n) = \beta p(n-1) + (1-\beta)[e_2(n) - e_1(n)]^2 \qquad (14)$$

is a rough (low-pass filtered) estimate of the power of the signal of interest. Selection of the forgetting factor, β, is rather easy. Typically, a choice of $\beta = 0.9$ ensures that $p(n)$ is adapted faster than any component filter. The overall structure, as given in Eqs. (13) and (14), is called normalized convex combination.

2.7 Normalized Fractional Least-Mean-Squares (NFLMS) Algorithm

The new idea is based on the fact that the normalized version of LMS algorithm has better performance than the standard LMS method. Furthermore, it has been shown that the FLMS algorithm, which is an improved version of the conventional LMS, has a faster convergence rate than LMS [11]. Thus, it is expected that using normalized version of FLMS (i.e., NFLMS) instead of FLMS leads to a better performance of adaptive filters. The update rule for NFLMS is:

$$\mathbf{w}(n+1) = \mathbf{w}(n) +$$

$$\left(\mu_1 + \mu_f \frac{\mathbf{w}^{1-v}(n)}{\Gamma(2-v)} \right) \frac{1}{\delta + \|\mathbf{u}(n)\|^2} e(n)\mathbf{u}(n) \qquad (15)$$

Here, v is the fractional order, μ_1 is the first order step-size, μ_f is the fractional order step-size, and $\delta > 0$. It has been shown that the performance of NFLMS is better than the standard LMS, NLMS, and FLMS algorithms [22].

3 Proposed Methods

In this section, our proposed methods, based on the fractional and/or normalized convex combination of fractional and/or normalized version of LMS component filters is explained.

3.1 Convex Combination of Normalized and/or Fractional Least-Mean-Squares Algorithm

One way of improving the performance of the whole convex combinational structure is to improve the performance of its component individual filters. In our previous work [23], we employed fractional LMS (i.e., FLMS) algorithm as individual filter. The result shows
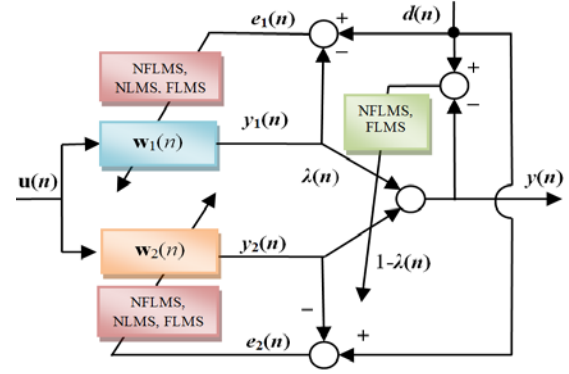
Fig. 3 The structure of the proposed convex combinational adaptive filters.

the superiority of the proposed algorithm. In this paper, we employ more LMS-based algorithms, such as NLMS and NFLMS as component filter in the structure of convex combination. Therefore, it is expected that using such algorithms as component filters leads to an increased convergence rate and reduced steady-state error of the overall filter. For this purpose, convex combinational adaptive filtering using NLMS, and NFLMS techniques in the implementation of the component filters is proposed. This is shown in Fig. 3.

3.2 Fractional Convex Combination of Adaptive Filters

As described in Section 2.5, the mixing parameter, $\lambda(n)$, is updated via the auxiliary parameter $a(n)$ (Eq. (10)), where $a(n)$ is updated in turn by the LMS algorithm (Eq. (12)). In order to improve the performance of mixing strategy in the convex combination, the update rule can be modified for $a(n)$ using fractional-based techniques. The proposed update relation for $a(n)$ in FLMS is as follows:

$$a(n+1) = a(n) +$$

$$\left(\mu_a + \mu_{af} \left(\frac{a^v(n)}{\Gamma(1-v)} \right) \mathrm{sgn}(a(n)) \right) \times \qquad (16)$$

$$\lambda(n)[1 - \lambda(n)]e(n)[e_2(n) - e_1(n)]$$

Here, μ_a and μ_{af} are the first order and the fractional order step-sizes, respectively.

To exploit the advantages of both normalization and fractional adaptation in the update rule for $a(n)$, the fractional normalized convex combination method is proposed. It will be shown that this idea leads to better performance of mixing parameter, $\lambda(n)$ The new update rule for $a(n)$ is given below:

$$a(n+1) = a(n) +$$

$$\left(\mu_a + \mu_{af} \left(\frac{a^v(n)\,\mathrm{sgn}(a^v(n))}{\Gamma(1-v)} \right) \right) \times \qquad (17)$$

$$\frac{\lambda(n)[1 - \lambda(n)]e(n)[e_2(n) - e_1(n)]}{p(n)}$$

where μ_a and μ_{af} are again the first order and the fractional order step-sizes, respectively, and

$$p(n) = \beta p(n-1) + (1-\beta)[e_2(n) - e_1(n)]^2 \qquad (18)$$

where β is the forgetting factor.

4 Evaluations and Experimental Results

For simulations, speech signals from the NOIZEUS database are used [24]. Noise signals are taken from the NOISEX-92 database [25]. The sampling rate of both speech and noise signals are set to 8000 Hz. Signals are digitized with 24 bit accuracy. The production of noisy speech follows two strategies. In the first strategy, a 30th-order FIR filter is used as the acoustic path to generate a random noisy signal, $d(n)$. In the second strategy, to simulate real conditions, the room impulse response given in [26] together with speech signal is used to generate the input noisy signal, $d(n)$. Fig. 4 illustrates the schematic diagram of the simulated room structure. Also, the corresponding impulse response is shown in Fig. 5.

To select an appropriate fractional order and fractional step-size for the simulations, learning curve of the FLMS algorithm is generated for different fractional orders and fractional step-sizes. For these simulations, input noisy signal obtained by the first strategy is used. From Fig. 6, it is observed that FLMS using fractional order 0.5 has the best performance. In addition, taking a fractional step-size, μ_f, equal to the first order step-size, μ_1, in the update rule of Eq. (7) appears to be the best choice among different simulations of the algorithm. Table 2 shows the parameters used in the implementation of algorithms.

In order to assess our proposed methods, the simulation results of twenty LMS-based algorithms using different subjective and objective criteria are investigated.

First, the performance of algorithms is studied by plotting their learning curves (i.e., MSE plots). For this purpose, a random white Gaussian noise with variance = -25 dB as clean input signal, $s(n)$, white Gaussian noise with mean=0 and variance = -55 dB as noise signal, $u(n)$, a 30th-order type I FIR filter as acoustic path ($L = 30$), and a 25th-order FIR filter as adaptive filter are used ($L = 25$).
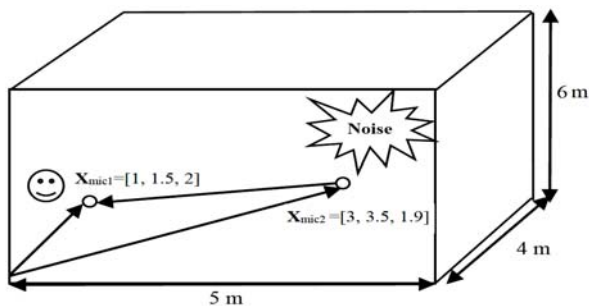
Fig. 4 Room structure used for the simulation of room impulse response. A room reverberation time of $RT_{60}= 0.4$ sec. is used for the simulations.
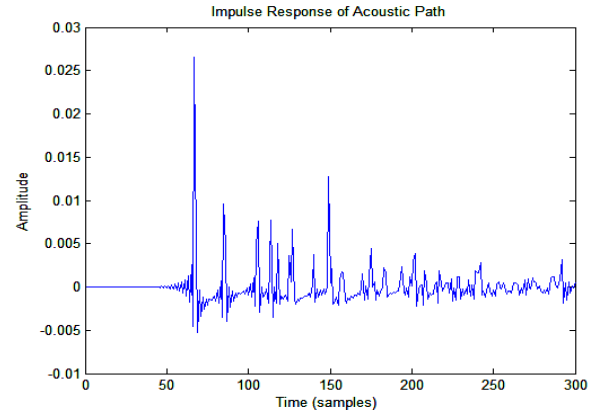

Fig. 5 The simulated room impulse response.

Table 2 The parameters used for implementation of algorithms.

Algorithms	Parameters	Range of Values
LMS, NLMS, FLMS, NFLMS	step-size (μ)	0.005
FLMS	step-size (μ)	0.005
	fractional step-size (μ_f)	0.005
	fractional derivation order (v)	0.5
Convex Combination	step-size for first filter (μ_1)	0.005
	step-size for second filter (μ_2)	0.001
	R	4
	μ_{max}	0.2
	B	0.95
	$a(0)$	0
	$\lambda(0)$	0.5
	μ_a	100
CC-FLMS, CC-NFLMS	fractional derivation order (v)	0.5
	fractional step-size of first component filter (μ_{f1})	0.005
	fractional step-size of second component filter (μ_{f2})	0.001
Normalized Convex Combination	$p(0)$	0.21
	μ_a	1
Fractional Convex Combination	μ_{af}	100
Fractional Normalized Convex Combination	μ_{af}	1

Fig. 7 shows the corresponding plots for the LMS, FLMS, NLMS, and NFLMS algorithms, obtained by averaging the results over 1000 runs. As the results of this simulation show, the proposed method (i.e., NFLMS) converges faster than other algorithms.

MSE PLOT for Different Fractional Order

(a)

MSE PLOT for Different Fractional Step Size

(b)
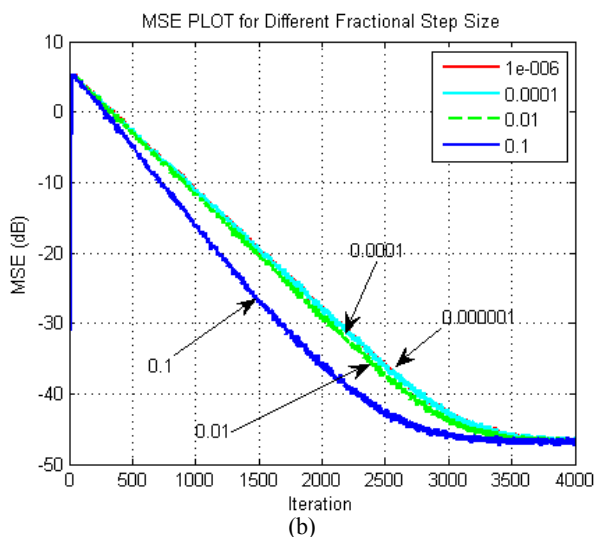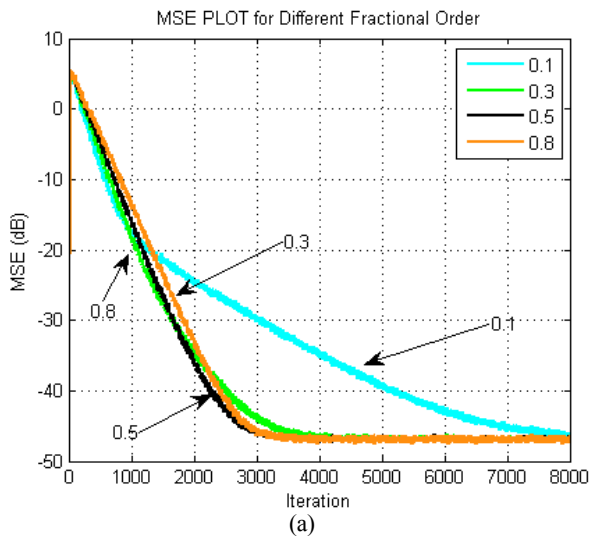
Fig. 6 The impact of fractional order (a) and fractional step-size (b) on the performance of the FLMS algorithm.
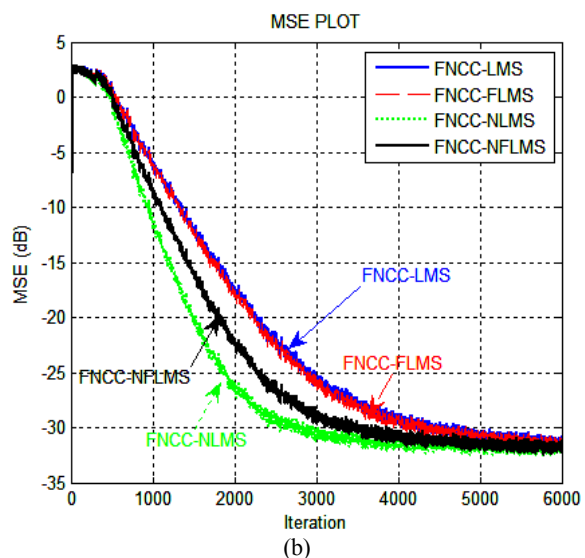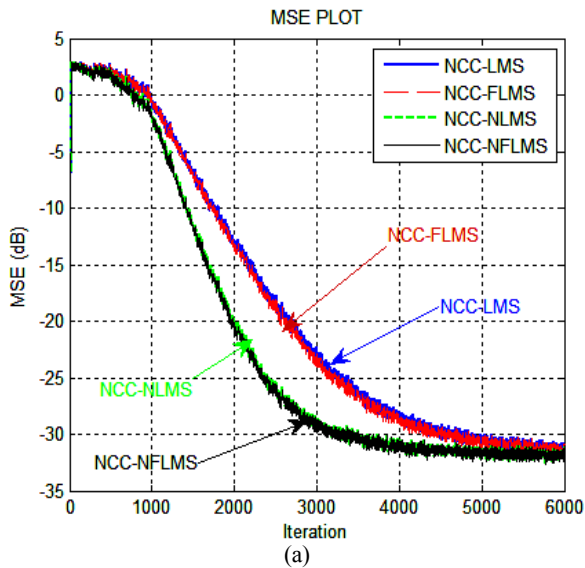

MSE PLOT
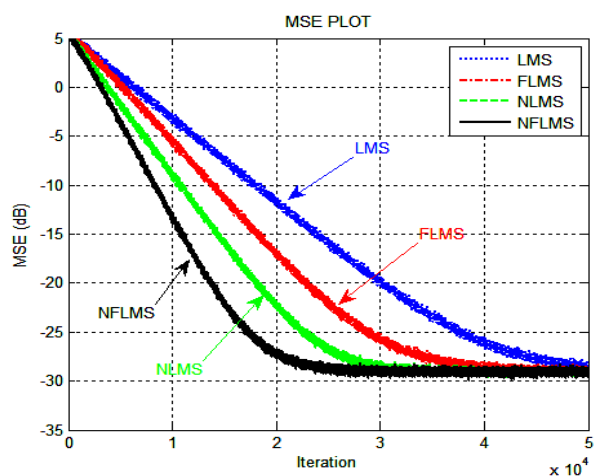
Fig. 7 The learning curves of LMS, FLMS, NLMS, and NFLMS for a 30th-order Type I FIR filter as acoustic path and 25th-order adaptive filter with a random signal as input clean signal and white noise as noise signal, averaged over 1000 runs.

To decide which type of component filters (i.e., LMS, FLMS, NLMS, and NFLMS) fits the most of the various convex combinational structures (i.e., Convex Combination (CC), Fractional Convex Combination (FCC), Normalized Convex Combination (NCC), and Fractional Normalized Convex Combination (FNCC)), the MSE plots drawn in Figs. 8 and 9 have been investigated. As it is observed from the simulated plots, NFLMS and NLMS have better convergence rates among all the convex combinational structures mentioned above. In general, it can be concluded that NLMS has the best performance in the sense of convergence rate among all simulated component filters.

Now, to decide which type of the mixing strategy (i.e., CC, FCC, NCC, and FNCC) fits the most with the best selected component filter (i.e., NLMS), the learning behavior of various convex combinational structures, shown in Fig. 10, have been studied. The results of this simulation show clearly that the FNCC mixing strategy gives the best performance among all the mentioned structures.


MSE PLOT

(a)

MSE PLOT

(b)
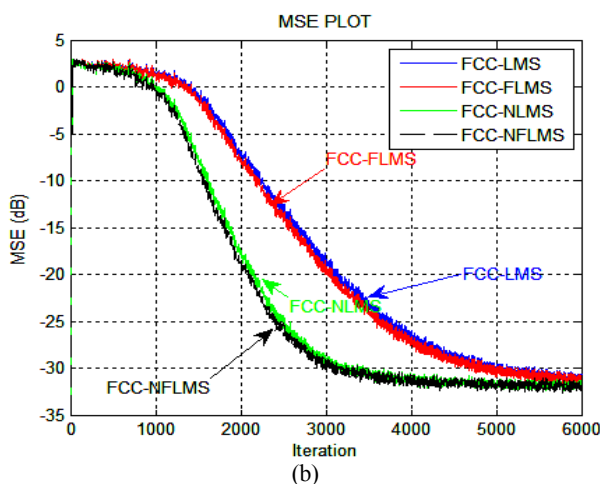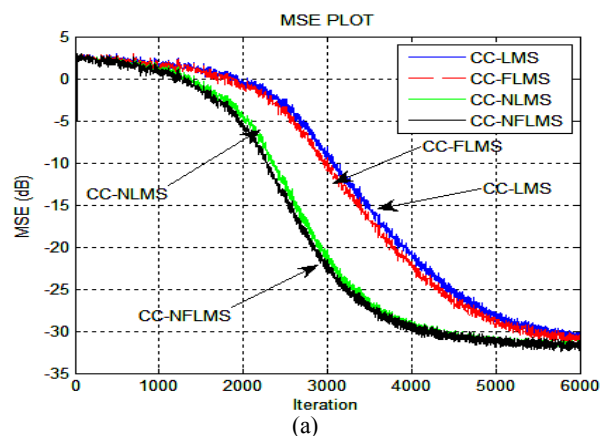
Fig. 8 The learning curves of CC-LMS, CC-FLMS, CC-NLMS, CC-NFLMS (a), and FCC-LMS, FCC-FLMS, FCC-NLMS, and FCC-NFLMS (b) for a 30th-order Type I FIR filter as acoustic path and 25th-order adaptive filter with a random signal as input clean signal and white noise as noise signal, averaged over 500 runs.

MSE PLOT

(a)

MSE PLOT

(b)

Fig. 9 The learning curves of NCC-LMS, NCC-FLMS, NCC-NLMS, NCC-NFLMS (a), and FNCC-LMS, FNCC-FLMS, FNCC-NLMS, and FNCC-NFLMS (b) for a 30th-order Type I FIR filter as acoustic path and 25th-order adaptive filter with a random signal as input clean signal and white noise as noise signal, averaged over 500 runs.

In the assessments of the proposed methods based on MSE plots, random signal is used as input clean signal (i.e. the first strategy).

Now, the performances of our proposed methods are examined in the case of real speech signals. The evaluations of the methods are conducted by inspecting the quality of the enhanced speech signal both in objective and subjective manners. In this part of simulations, the room impulse response is used to simulate real acoustic conditions. As noise signal, babble noise with SNRs of 0 dB and 10 dB, and car noise with SNRs of -5 dB and 5 dB are used.
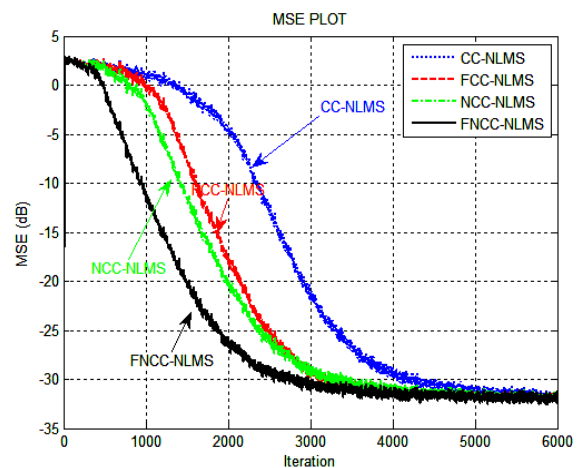
Fig. 10 The learning curves of CC-NLMS, FCC-NLMS, NCC-NLMS, and FNCC-NLMS for a 30th-order Type I FIR filter as acoustic path and 25th-order adaptive filter with a random signal as input clean signal and white noise as noise signal, averaged over 500 runs.

As objective evaluation criteria, the segmental SNR and PESQ tests [27-28] are employed. The results are shown in Figs. 11, 12, 13, and 14 for different noise sources and different input SNR values. As it can be seen from the figures, the speech signal enhanced by FNCC-NLMS has the best quality, compared with that obtained from other methods. This is in accordance with the MSE evaluation results obtained by using a random clean signal.

In order to assess the proposed algorithms subjectively, the MUlti Stimulus test with Hidden Reference and Anchor (MUSHRA) is used, which is an ITU-R Recommendation BS.1534-1 [29] as implemented in [30, 31]. This method has been used in the framework of speech separation problem to assess the quality of separated speech signal [32].

Before beginning the listening tests, listeners are informed about the aim of the listening task, namely, the assessment of speech quality. For this purpose, listeners are asked to pay attention to the amount of background noise and speech distortion. Here, a training phase is conducted to make listeners familiar with the test procedure. First, subjects are allowed to listen to the test speech signals without evaluating them. Then, they are asked to give scores to the processed signals.

The subjects (i.e., human listeners) are provided with test utterances plus one reference and one hidden anchor, and are asked to rate different signals on a scale of 0 to 100, where 100 represents the best score. The listeners are permitted to listen to each sentence several times and always have access to clean signal reference.

The test signals are the same as those used for the objective evaluation. Two types of noises (i.e., Car noise and Babble noise) are used during the listening tests. A total of 10 listeners (3 females and 7 males between the ages of 18 and 30) have participated in these tests.
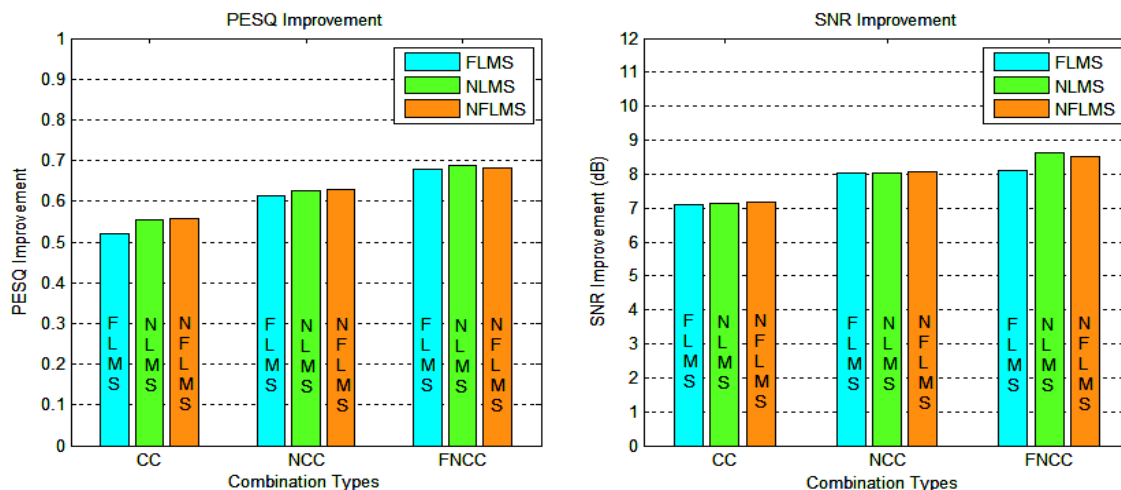
Fig. 11 PESQ and SNR improvements for the CC-FLMS, CC-NLMS, CC-NFLMS, NCC-FLMS, NCC-NLMS, NCC-NFLMS, FNCC-FLMS, FNCC-NLMS, and FNCC-NFLMS algorithms obtained by using a real speech signal as input clean signal and babble noise with SNR of 0 dB as input noise signal.
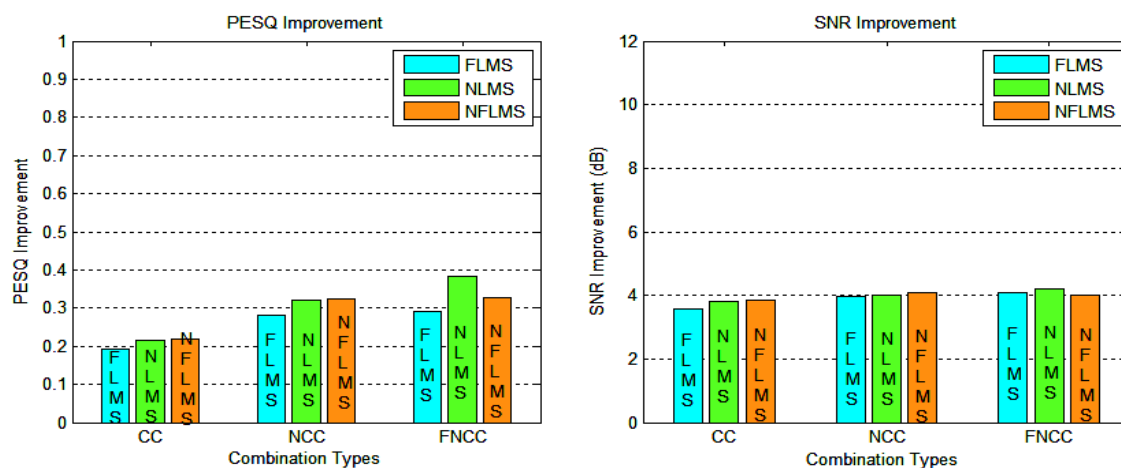
Fig. 12 PESQ and SNR improvements for the CC-FLMS, CC-NLMS, CC-NFLMS, NCC-FLMS, NCC-NLMS, NCC-NFLMS, FNCC-FLMS, FNCC-NLMS, and FNCC-NFLMS algorithms obtained by using a real speech signal as input clean signal and babble noise with SNR of 10 dB as input noise signal.
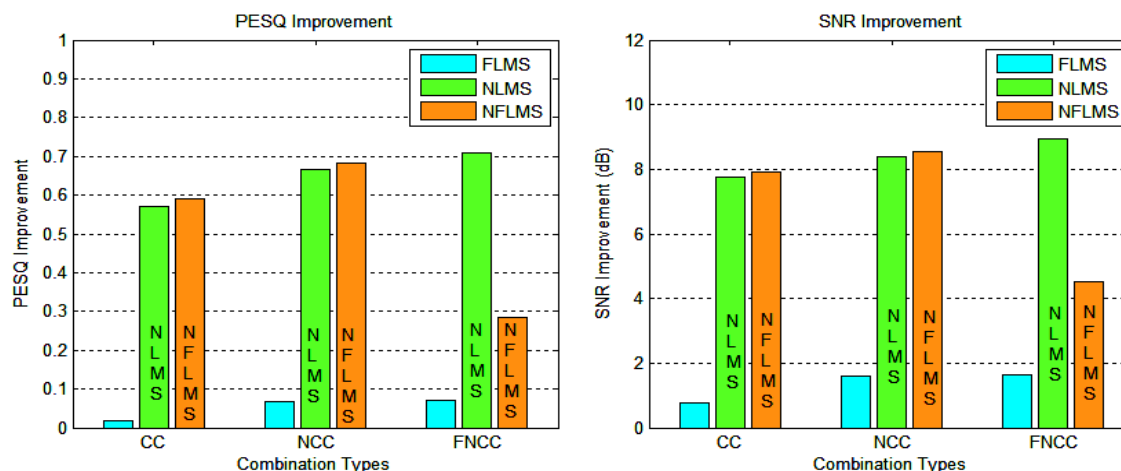
Fig. 13 PESQ and SNR improvements for the CC-FLMS, CC-NLMS, CC-NFLMS, NCC-FLMS, NCC-NLMS, NCC-NFLMS, FNCC-FLMS, FNCC-NLMS, and FNCC-NFLMS algorithms obtained by using a real speech signal as input clean signal and car noise with SNR of -5 dB as input noise signal.
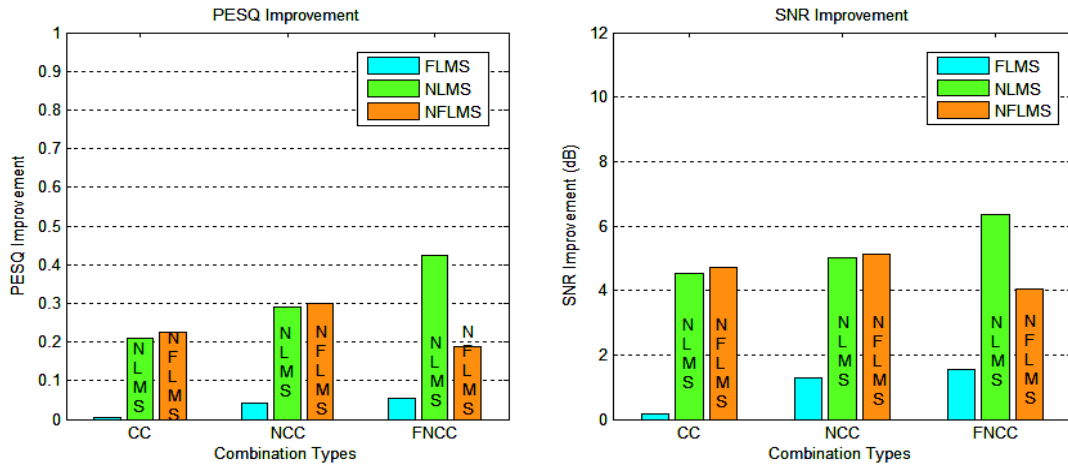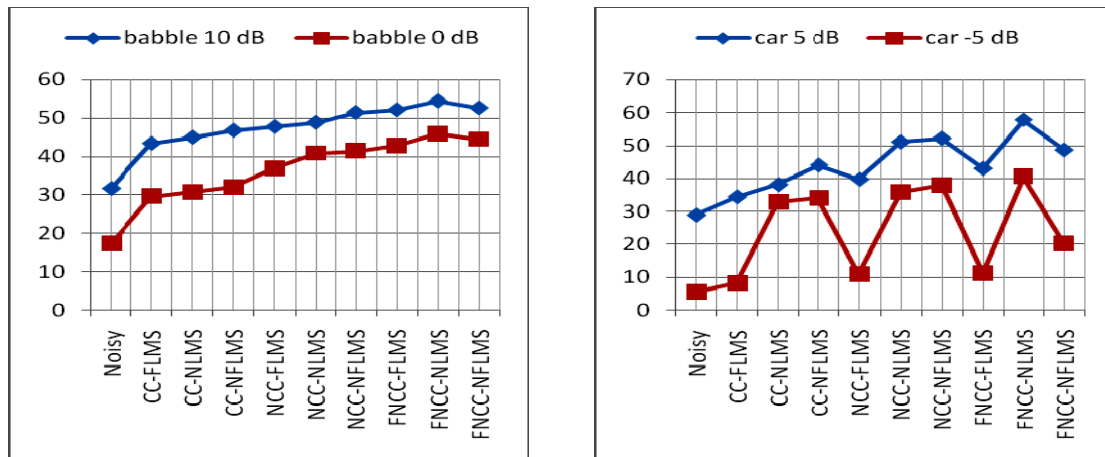
Fig. 14 PESQ and SNR improvements for the CC-FLMS, CC-NLMS, CC-NFLMS, NCC-FLMS, NCC-NLMS, NCC-NFLMS, FNCC-FLMS, FNCC-NLMS, and FNCC-NFLMS algorithms obtained by using a real speech signal as input clean signal and car noise with SNR of 5 dB as input noise signal.

Fig. 15 The MUSHRA listening test results obtained by using a real speech signal as input clean signal, and babble noise with SNRs of 0 dB and 10 dB (left panel) and car noise with SNRs of -5 dB and 5 dB (right panel) as input noise signals.

Fig. 15 shows the results of subjective listening tests for each algorithm and different noise types. By examining the results of listening tests, it is obvious that the FNCC-NLMS method produces the highest speech quality in speech enhancement system, as compared with other simulated algorithms. The superior performance of the FNCC-NLMS method is in agreement with the results obtained during the objective evaluation tests, and is again in accordance with the MSE learning curves obtained by random clean signal.

5 Conclusions

In this paper, new convex combinational adaptive filtering methods are proposed in the framework of speech enhancement system. The proposed methods utilize the idea of normalization and fractional derivative, both in the design of different convex mixing strategies and in their related component filters.

To evaluate the performance of this new idea, in the first strategy, the simulations of learning curves (i.e., MSE plots) are examined using random signal instead of clean speech signal. As it can be inferred from the behaviors of the MSE plots, it can be verified that the idea of normalization and fractional derivative leads to improved performance in the sense of convergence rate in the whole structure of convex combinational adaptive filtering. The study of MSE learning curves shows clearly that the FNCC-NLMS algorithm has the best performance among all the proposed (i.e., CC-NLMS, CC-NFLMS, FCC-LMS, FCC-NLMS, FCC-FLMS, FCC-NFLMS, NCC-FLMS, NCC-NFLMS, FNCC-LMS, FNCC-FLMS, FNCC-NLMS, and FNCC-NFLMS) and simulated algorithms.

In the second strategy, a real input speech signal is used in the simulations and the quality of enhanced speech is investigated, both objectively and subjectively.

To this aim, FNCC-NLMS, as selected by the MSE evaluations, is compared with other convex combinational methods. As objective evaluation, SNR and PESQ improvements, obtained from different methods, are compared. From the results, it can be concluded that the speech enhanced by FNCC-NLMS has the highest quality.

To assess the performance of FNCC-NLMS subjectively, listening tests have been conducted for the enhanced (real) speech obtained by applying the same methods as used in the objective evaluations (i.e., SNR and PESQ tests). The results show, once again, that the speech signal enhanced by FNCC-NLMS presents the highest quality among the signals obtained by all simulated methods.

In general, the powerful aspects of our proposed methods can be stated to be their low complexity, as expected with all LMS-based methods, together with their high convergence rate.

As future work, the new adaptive filtering structures can be incorporated in other adaptive signal processing applications.

Appendix

The computational complexity of the most important relations used in different adaptive algorithms are shown in Table 3. The computations have been performed by considering the number of additions and multiplications in each iteration assuming that the length of filter is L. The interpretation of the results of this table confirm the fact that the computational load depends both on the number of operations (i.e., additions and multiplications) and the use of nonlinearities, such as sign function, sigmoidal function, and fractional order, used in the update rules. Also, it can be observed that the computational burden is remarkably increased by using convex combination structure in speech enhancement systems. It can generally be concluded that the overall computational load of the convex combinational adaptive filtering is almost twice that of traditional algorithms.

Table 3 The computational complexity of important relations.

Equation	No. of Additions	No. of Multiplications
Eq. (1)	L	L
Eq. (3)	L	$2L$
Eq. (4)	L	$5L$
Eq. (8)	2L	$6L$
Eq. (9)	L	$2L$
Eq. (12)	$3L$	$5L$
Eq. (13)	$3L$	$6L$
Eq. (14)	$3L$	$4L$
Eq. (16)	$3L$	$6L$
Eq. (17)	$3L$	$9L$
Eq. (18)	L	$4L$

References

[1] P. C. Loizou, *Speech Enhancement Theory and Practice*, CRC Press, 1st Edition, 2007.

[2] M. Feder, A. V. Oppenheim and E. Weistein, "Maximum Likelihood Noise Cancellation Using the EM Algorithm", *IEEE Trans. on Acoustics, Speech and Signal Processing*, Vol. 37, No. 2, pp. 204-216, 1989.

[3] B. Widrow and M. E. Hoff, "Adaptive Switching Circuits", *Proc. IRE WESCON Conv. Rec. 4*, pp. 96-104, 1960.

[4] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1985.

[5] S. Haykin, *Adaptive Filter Theory*, 4th Edition. Upper Saddle River, NJ: Prentice-Hall, 2002.

[6] A. H. Sayed, *Fundamentals of Adaptive Filtering*, John Wiley & Sons, New York, 2003.

[7] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, John Wiley & Sons, New York, 1993.

[8] S. G. Samko, A. A. Kilbas and O. I. Marichev, *Fractional Integrals and Derivatives: Theory and Applications*, Reprint Taylor & Francis Books Ltd, London, 2002.

[9] C. C. Tseng, "Design and Application of Variable Fractional Order Differentiator", *IEEE Asia-Pacific Conference on Circuits and System,* Vol. 1, pp. 405-408, Taivan, Dec. 2004.

[10] C. C. Tseng, "Design of Fractional Order Digital FIR Differentiators", *IEEE Signal Processing Letter*, Vol. 8, No. 3, pp. 77-79, 2001.

[11] R. Muhammad Asif Zahoor, "A Modified Least-Mean-Squares Algorithm Using Fractional Derivative and its Application to System Identification", *European Journal of Scientific Research*, ISSN: 1450-216, Vol. 35, No. 1, pp 14-21, 2009.

[12] B. Shoaib and I. M. Qureshi, "A Modified Fractional Least-Mean-Square Algorithm for Chaotic and Nonstationary Time Series Prediction", *Chinese Physics,* Vol. 23, No. 3, pp. 305-302, 2014.

[13] D. S. Kumar and N. K. Rout, "FLMS Algorithm for Acoustic Echo Cancellation and Its Comparison with LMS", In *1st International Conference on Recent Advances in Information Technology (RAIT),* pp. 852-856, March 2012.

[14] J. Li, D. Feng and Y. Guowei, "Maximum Likelihood Least-Squares Identification Method for Input Nonlinear Finite Impulse Response Moving Average Systems", *Mathematical and Computer Modelling,* Vol. 55, No. 3, pp. 442-450, 2012.

[15] S. S. Kozat, A. T. Erdogan, A. C. Singer and A. H. Sayed, "Steady-state MSE Performance Analysis of Mixture Approaches to Adaptive

Filtering", *IEEE Transactions on Signal Processing*, Vol. 58, No. 8, pp. 4050-4063, 2010.

[16] N. J. Bershad, J. C. M. Bermudez and J.-Y. Tourneret, "An Affine Combination of Two LMS Adaptive Filters Transient Mean-Square Analysis", *IEEE Transactions on Signal Processing*, Vol. 56, No. 5, pp. 1853-1864, 2008.

[17] J. Arenas-García, A. l. 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, March 2006.

[18] A. Abdollahy and M. Geravanchizadeh, "Speech Enhancement Using Combinational Adaptive Filtering", *4ᵗʰ Int. Symp. on Telecommunications* (*IST*), Tehran, pp. 774-776, 2008.

[19] M. Ferrer, A. Gonzalez, M. de Diego and G. Pinero, "Convex Combination Filtered-x Algorithms for Active Noise Control Systems", *IEEE Transactions on Audio, Speech, and Language Processing,* Vol. 21, No. 1, pp. 156-167, 2013.

[20] L. A. Azpicueta-Ruiz, A. R. Figueiras-Vidal, J. Arenas-Garcia, "A Normalized Adaptation Scheme for the Convex Combination of Two Adaptive Filters", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, (*ICASSP*), Las Vegas, pp.3301-3304, 2008.

[21] M. Weilbeer, *Efficient Numerical Methods for Fractional Differential Equations and Their Analytical Background*, Ph.D. thesis Papierflieger, Taschenbuch, 2006.

[22] M. Geravanchizadeh and S. G. Osgouei, "Speech Enhancement Using Normalized Fractional Least-mean-squares Algorithm", *19ᵗʰ Iranian Conf. on Electrical Engineering* (*ICEE*), Tehran, pp. 1-5, 2011.

[23] S. G. Osgouei and M. Geravanchizadeh, "Speech Enhancement Using Convex Combination of Fractional Least-mean-squares Algorithm", *5ᵗʰ International Symposium on Telecommunications* (*IST*), Tehran, pp. 869-872, 2010.

[24] http://www.utdallas.edu/~loizou/speech/noizeus, Available: (Nov.22.2011).

[25] http://www.speech.cs.cmu.edu/comp.speech/Section1/Data/noisex.html, Available: (Nov. 24.2011).

[26] E .A. P. Habets**,** "Room Impulse Response Generator", *Technische Universiteit Eindhoven, Tech. Rep.* 2, No. 2.4, 2006.

[27] A. W. Rix, J. G. Beerends, M. P. Hollier and A. P. Hekstra, "Perceptual Evaluation of Speech Quality (PESQ), A New Method for Speech Quality Assessment of Telephone Networks and Codecs", *Proceeding of 26ᵗʰ IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, Vol. 2, pp. 749-752, Utah, USA, 7-11 May 2001.

[28] Y. Hu and P. C. Loizou, "Evaluation of Objective Measures for Speech Enhancement", *Proceedings of Interspeech*, Philadelphia, PA, 2006.

[29] ITU-R, "Recommendation BS1543-1: Method for the Subjective Assessment of Intermediate Quality Level of Coding Systems", 2001.

[30] M. Hecker and C. Williams, "Choice of Reference Conditions for Speech Preference Tests", *J. Acoust. Soc. Am.*, Vol. 39, No. 5A, pp. 946-952, 1966.

[31] E. Vincent, "MUSHRAM: A MATLAB Interface for MUSHRA Listening Tests," [Online] available (Aug.1.2013): http://www.elec.qmul.ac.uk/people/emmanuelv/mushram.

[32] P. Mowlaee, C. G. Mads and J. H. Søren, "New Results on Single-Channel Speech Separation Using Sinusoidal Modeling", *IEEE Transactions on Audio, Speech, and Language Processing,* Vol. 5, No. 19, pp.1265-1277, 2011.

Sina Ghalamiosgouei was born in Tabriz, Iran, in 1987. He received the B.Sc. and M.Sc. degrees both in Communication Engineering from the University of Tabriz in 2009 and 2011, respectively. He is currently pursuing Ph.D. degree in Faculty of Computer Eng. at University of Tabriz, Tabriz, Iran. His current research interests are focused on Speech Enhancement, Binaural Signal Processing, Adaptive Signal Processing, Fractional Signal Processing, and Pattern Classification.

Masoud Geravanchizadeh received the B.Sc. degree in Electronics Engineering from the University of Tabriz, Tabriz, Iran, in 1986, and the M.Sc. and Ph.D. degrees in Signal Processing from the Ruhr-University Bochum, Bochum, Germany, in 1995 and 2001, respectively. Since 2005, he has been with the Faculty of Electrical and Computer Eng., at the University of Tabriz, Tabriz, where he is currently an Assistant Professor. His research interests include Speech Enhancement, Sound Source Localization, Speech Separation, Stochastic Signal Processing, and Binaural Signal Processing.