# Novel Sparse Algorithms Based on Lyapunov Stability for Adaptive System Identification

*Rakesh POGULA [1], T. Kishore KUMAR [1], Felix ALBU [2]*

[1] Dept. of Electronics and Communication Engineering, National Inst. of Technology Warangal, 506004 Warangal, India
[2] Dept. of Electronics, Valahia University of Targoviste, 130082 Targoviste, Romania

rakesh.pogula453@ gmail.com,  kishorefr@gmail.com,  felix.albu@valahia.ro

**Abstract.** *Adaptive filters are extensively used in the identification of an unknown system. Unlike several gradient-search based adaptive filtering techniques, the Lyapunov Theory-based Adaptive Filter offers improved convergence and stability. When the system is described by a sparse model, the performance of Lyapunov Adaptive (LA) filter is degraded since it fails to exploit the system sparsity. In this paper, the Zero-Attracting Lyapunov Adaptation algorithm (ZA-LA), the Reweighted Zero-Attracting Lyapunov Adaptation algorithm (RZA-LA) and an affine combination scheme of the LA and proposed ZA-LA filters are proposed. The ZA-LA algorithm is based on $\ell_1$-norm relaxation while the RZA-LA algorithm uses a log-sum penalty to accelerate convergence when identifying sparse systems. It is shown by simulations that the proposed algorithms can achieve better convergence than the existing LMS/LA filter for a sparse system, while the affine combination scheme is robust in identifying systems with variable sparsity.*

## Keywords

Sparse system identification, Lyapunov adaptive filter (LA), $\ell_1$-norm, Zero-Attracting LA (ZA-LA), Reweighted ZA-LA (RZA-LA), affine combination, convergence, mean square deviation, mean square error

## 1. Introduction

Adaptive filtering algorithms play a significant role in system identification applications, e.g., channel estimation and echo cancellation [1]. Ideally, the adaptive filter with high convergence rate, stability, good tracking capability and robustness to random noise is desirable for many applications. The widely used optimization technique for optimal filter design is the gradient descent method. Least Mean Square (LMS) algorithm developed by Widrow and Hoff [2] is the most widely used adaptive algorithm due to its computational simplicity and proven robustness. Lyapunov Adaptive Filtering (LA) algorithms in the sense of the Lyapunov stability theory were proposed in [3–6] to overcome the problems faced by gradient descent-based techniques such as slow rate of convergence, sensitivity to variations in the eigenvalue spread and local minima problem. Moreover, the LA algorithm is independent of the stochastic properties of the input signal and additive noise. However, the LA algorithm suffers from poor convergence performance when the underlying system to be identified is sparse such as network and acoustic echo path [7], digital TV transmission channel [8], and underwater channel [9]. In general, the sparse FIR system is characterized by its impulse response which consists of very few active coefficients among many inactive ones [10–12].

Conventional adaptive algorithms neglect the sparse information which is present in the system that leads to degrade their performance when estimating the sparse channels. Recent studies on system identification specify that by utilizing the *a priori* knowledge about the system sparsity, the estimation performance can be improved substantially. This motivated the researchers towards developing sparse adaptive filtering algorithms in the last few years. Inspired by the theory of Compressed Sensing (CS) [13], [14] and LASSO [15], $\ell_1$-norm relaxation is utilized to exploit system sparsity. In [16] sparse adaptive filtering approach based on LMS algorithm is developed by incorporating $\ell_1$-norm penalty into its cost function which is termed as Zero-Attracting LMS (ZA-LMS). This approach is easy to implement and performs well when the system is highly sparse, but it fails as the system sparsity decreases. Based on reweighted $\ell_1$-minimization sparse recovery algorithm [17], an improved version of ZA-LMS i.e. Reweighted ZA-LMS (RZA-LMS) is proposed in [16]. The RZA-LMS algorithm performs better in less sparse conditions but at the cost of increased complexity. Following these ideas, in this paper, we propose two sparse Lyapunov Adaptation algorithms namely, Zero-Attracting Lyapunov Adaptation algorithm (ZA-LA) and Reweighted Zero-Attracting Lyapunov Adaptation algorithm (RZA-LA).

Over the past decade or so, a combination of adaptive filters has proven to be an efficient way to handle systems with variable sparsity. In [18], [19], an adaptive convex combination of two LMS filters with different parameter setting is proposed to alleviate the speed of convergence vs

the residual error trade-off. In [20], a mixture approach of adaptively combining LMS and ZA-LMS algorithm using a convex combination has been proposed to achieve robustness against time-varying system sparsity. This approach is extended [21] to colored (correlated) input signal using a convex combination of the Affine Projection Algorithm (APA) and Zero Attracting Affine Projection Algorithm (ZA-APA), and in [22] steady state mean square analysis of convex combination in the context of acoustic echo cancellation is performed. In [23], a convex combination approach using another variant of sparse adaptive filtering, i.e. Improved Proportionate Normalized LMS (IPNLMS) algorithm is proposed. In all these papers, the authors have used the convex combination to combine effectively the outputs of the individual adaptive filters. The affine combination as a generalization of the convex combination is studied in [24], and in [25] affine combination analysis was extended for colored inputs and nonstationary environments. In [26], transient analysis for the affine combination of two NLMS adaptive filters is studied. In [27], [28], it is demonstrated that affine combination results in faster convergence than the convex combination of two adaptive filters. So, in this paper, we use affine combination scheme to combine LA and ZA-LA filters to handle systems with variable sparsity. To the best of our knowledge, no paper has reported on system identification using Lyapunov adaptive filter algorithm with sparsity constraints to identify an unknown sparse system.

The rest of the paper is organized as follows. Section 2 describes the system model and reviews the LA algorithms. In Sec. 3, we propose two adaptive sparse LA algorithms. In Sec. 4 we compare the algorithms in terms of computational complexity. Section 5 highlights the affine combination scheme and its stability is analyzed in Sec. 6. Simulation results of the proposed LA algorithms are presented in Sec. 7. Finally, we conclude the paper in Sec. 8.

## 2. Review of LA Algorithms

The basic idea of sparse system identification is to improve the filtering/estimation performance by utilizing the inherent sparse structure information. The block diagram of the sparse system identification is shown in Fig. 1. We consider an $N$ length FIR filter coefficient vector $\mathbf{w}_0 = [w_0, w_1, \ldots, w_{N-1}]^T$ and the input signal vector $\bar{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^T$. The input $\bar{x}(n)$ is applied to both the adaptive filter and the unknown sparse system. The desired signal $d(n)$ that is generally corrupted by the observation noise $v(n)$ is given by $d(n) = w_0^T \bar{x}(n) + v(n)$. The output estimate $y(n)$ of the adaptive filter $\overline{W}(n)$ is subtracted from the reference signal $d(n)$ to produce an error signal $e(n)$ [29]. The error signal $e(n)$ is then used by the adaptive algorithm iteratively to manipulate the filter coefficients such that the error is minimized.

The cost function $V(n)$ of the LA filter is defined as the square of error between the desired reference input and
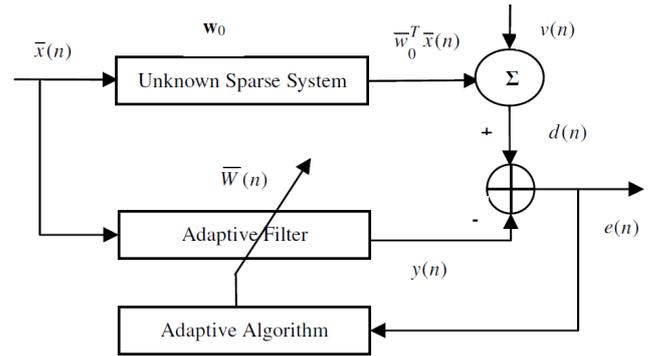


**Fig. 1.** Block diagram of sparse system identification.

filtered output which is considered as the Lyapunov function [3], [4],

$$V(n) = e^2(n) \tag{1}$$

where

$$e(n) = d(n) - y(n). \tag{2}$$

The Lyapunov adaptive filter weight update law is properly chosen such that $\Delta V(k) = V(k) - V(k-1)$ is negative definite at each iteration. According to Lyapunov stability theory [30], when $\Delta V(k) < 0$, the output of the adaptive filter can asymptotically converge to the desired signal which means that the error $e(n) \to 0$.

For the given desired signal $d(n)$ and filter output $y(n) = \overline{W}^T(n)\bar{x}(n)$ the LA weight update rule is as follows:

$$\overline{W}(n) = \overline{W}(n-1) + g(n)\alpha(n) \tag{3}$$

and

$$g(n) = \frac{\bar{x}(n)}{\lambda_1 + \left\|\bar{x}(n)\right\|^2}\left[1 - \kappa\frac{|e(n-1)|}{\lambda_2 + |\alpha(n)|}\right] \tag{4}$$

where $g(n)$ is the adaptation gain and $\alpha(n)$ is *a priori* estimation error defined as

$$\alpha(n) = d(n) - \overline{W}^T(n-1)\bar{x}(n), \tag{5}$$

$0 \leq \kappa < 1$ and $\lambda_1$, $\lambda_2$ are small positive constants to prevent the singularities in the adaptation gain. The parameter $\kappa$ controls the convergence rate of the algorithm.

To improve the LA filter tracking performance and robustness, the Lyapunov function is redefined as $V(n) = \beta^n e^2(n)$ and the adaptation gain $g(n)$ of the Lyapunov adaptive algorithm is modified [5], [6] as

$$g(n) = \frac{\bar{x}(n)}{\lambda_1 + \left\|\bar{x}(n)\right\|^2}\left[1 - \frac{|e(n-1)|}{\lambda_2 + |\alpha(n)|\beta^{n/2}}\right] \tag{6}$$

where $\beta > 1$ is a constant parameter.

To achieve faster convergence rate and lower steady state error performance in the case of noisy environments, a step size parameter, $\mu$ is included in the adaptation gain which is given by [31]

$$g(n) = \frac{\mu . \bar{x}(n)}{\lambda_1 + \left\| \bar{x}(n) \right\|^2} \left[ 1 - \kappa \frac{|e(n-1)|}{\lambda_2 + |\alpha(n)|} \right]. \qquad (7)$$

The adaptation gain $g(n)$ for the Lyapunov adaptive algorithm can also be defined as

$$g(n) = \frac{\bar{x}(n)}{\lambda_1 + \left\| \bar{x}(n) \right\|^2} \left[ 1 - \frac{|e(n-1)|}{\lambda_2 + \beta^{n/2}(n-1)|\alpha(n)|} \right] \qquad (8)$$

where the adaptive adaptation gain rate is adjustable in order to improve the tracking capability of the algorithm [32–34]

$$\beta(n) = 1 + \frac{e^2(n-1)}{e^2(n)}. \qquad (9)$$

# 3. Proposed Sparse LA Algorithms

## 3.1 Zero-Attracting Lyapunov Adaptation Algorithm (ZA-LA)

The cost function $V_1(n)$ of ZA-LA is defined by combining the Lyapunov function of LA filter with the $\ell_1$-norm penalty of the weight vector

$$V_1(n) = \underbrace{\beta^n e^2(n)}_{V(n)} + \gamma_{ZA} \underbrace{\left\| \bar{W}(n) \right\|_1}_{l_1 - \text{norm}} \qquad (10)$$

where $\gamma_{ZA} > 0$ denotes a regularization parameter which balances the error term and system sparsity.

We then have

$$\Delta V_1(n) = V_1(n) - V_1(n-1)$$
$$= \beta^n e^2(n) - \beta^{n-1} e^2(n-1) + \gamma_{ZA} (\left\| \bar{W}(n) \right\|_1 - \left\| \bar{W}(n-1) \right\|_1)$$
$$= \beta^n [d(n) - \bar{W}^T(n)\bar{x}(n)]^2 - \beta^{n-1} e^2(n-1) + \gamma_{ZA} (\left\| \bar{W}(n) \right\|_1 - \left\| \bar{W}(n-1) \right\|_1)$$
$$= \beta^n [d(n) - (W^T(n-1) + g^T(n)\alpha(n))\bar{x}(n)]^2 - \beta^{n-1} e^2(n-1) + \gamma_{ZA} (\left\| \bar{W}(n) \right\|_1 - \left\| \bar{W}(n-1) \right\|_1)$$
$$= \beta^n [\alpha(n) - \alpha(n)g^T(n)\bar{x}(n)]^2 - \beta^{n-1} e^2(n-1) + \gamma_{ZA} (\left\| \bar{W}(n) \right\|_1 - \left\| \bar{W}(n-1) \right\|_1). \qquad (11)$$

Using the adaptation gain $g(n)$ given in (6), we obtain

$$\Delta V_1(n) = -(\beta^{n-1} - 1)e^2(n-1) + \gamma_{ZA}(\left\| \bar{W}(n) \right\|_1 - \left\| \bar{W}(n-1) \right\|_1). \qquad (12)$$

The ZA-LA filter update rule is given by

$$\bar{W}(n) = \bar{W}(n-1) + g(n)\alpha(n) - \gamma_{ZA} \operatorname{sgn}(\bar{W}(n)) \qquad (13)$$

where $\operatorname{sgn}(\cdot)$ is the sign function. The ZA-LA algorithm complexity is slightly higher than that of LA algorithm due to the third term of (13).

## 3.2 Reweighted Zero-Attracting Lyapunov Adaptation Algorithm (RZA-LA)

The RZA-LA cost function is represented by

$$V_2(n) = \beta^n e^2(n) + \gamma_{RZA} \sum_{i=1}^N \log(1 + \varepsilon |w_i(n)|) \qquad (14)$$

where $\gamma_{RZA} > 0$ is a regularization parameter which balances the estimation error term and sparsity of $\bar{W}(n)$.

According to the log–det heuristic approach used in [35], [36], the zero attractor term in (14) yields to a convex optimization problem. To establish this connection, consider the problem

$$\text{minimize} \sum_i \log(1 + \varepsilon |w_i|)$$
$$\text{subject to } \bar{W} \in R^n. \qquad (15)$$

Iterative linearization of this objective function gives

$$\bar{W}(n+1) = \arg\min \sum_i \frac{|w_i|}{1 + \varepsilon |w_i(n)|}. \qquad (16)$$

If $w_i(n)$ is small, its weighting factor in the next minimization step, $(1 + \varepsilon |w_i(n)|)^{-1}$, is large. So the small entries in $\bar{W}(n)$ are pushed towards zero. Thus, the log-sum penalty function has the potential to be much more sparsity encouraging than the $\ell_1$-norm.

The RZA-LA filter update is defined as

$$\bar{W}(n) = \bar{W}(n-1) + g(n)\alpha(n) - \gamma_{RZA} \frac{\operatorname{sgn}(\bar{W}(n))}{(1 + \varepsilon_{RZA} |\bar{W}(n)|)}. \qquad (17)$$

The RZA-LA algorithm complexity is slightly higher than that of LA algorithm due to the third term of (17).

# 4. Computational Complexity

The numerical complexity in terms of additions, multiplications, and divisions of the proposed sparse adaptive algorithms and those of the LMS algorithm and its sparse variants are shown in Tab. 1.

| Algorithms | Addition | Multiplication | Division |
|---|---|---|---|
| LMS | $2N$ | $2N + 1$ | - |
| ZA-LMS | $3N$ | $3N + 1$ | - |
| RZA-LMS | $3N + 1$ | $3N + 2$ | $N$ |
| LA | $3N + 5$ | $3N + 3$ | $N + 1$ |
| ZA-LA | $4N + 5$ | $4N + 3$ | $N + 1$ |
| RZA-LA | $4N + 6$ | $4N + 4$ | $2N + 1$ |

**Tab. 1.** Comparison of computational complexity of the investigated algorithms.

It can be seen from Tab. 1, that the proposed sparse algorithms have a moderate computational complexity increase when compared with the original algorithms.

# 5. Affine Combination of LA and ZA-LA Algorithms

In order to handle the system with varying level of sparseness, we have proposed to combine LA and ZA-LA algorithm using an affine combination approach. The configuration of the proposed affine combination scheme is shown in Fig. 2 in which Filter 1 is updated using LA algorithm (3), and Filter 2 is updated using ZA-LA algorithm as given in (13), respectively. We will hereafter call this filter as the Affine Combined Lyapunov Adaptation (ACLA) filter.

The output signal of ACLA filter is given by

$$y(n) = \lambda(n)y_1(n) + [1 - \lambda(n)]y_2(n) \qquad (18)$$

where $\lambda(n)$ is the mixing parameter and can be any real number, and $y_1(n)$ and $y_2(n)$ denotes the output of the individual filters, i.e. $y_i(n) = \overline{W}_i^T(n-1)\,\overline{x}(n)$, $i = 1,2$. The a priori error signal $e_a(n)$ is obtained by subtracting the output signal of ACLA filter from the output signal of the unknown system,

$$e_a(n) = y_0(n) - y(n) =$$
$$y_0(n) - \lambda(n)y_1(n) - [1 - \lambda(n)]y_2(n) \qquad (19)$$

where

$$y_0(n) = \overline{w}_0^T\,\overline{x}(n) = d(n) - v(n) \qquad (20)$$

To find the mixing parameter $\lambda(n)$, the derivative of mean square of the a priori error with respect to $\lambda(n)$ is evaluated and equated to zero.

$$\frac{\partial E[|e_a(n)|^2]}{\partial \lambda(n)} =$$
$$2E[(y_0(n) - \lambda(n)y_1(n) - (1 - \lambda(n))y_2(n))(-y_1(n) + y_2(n))]$$
$$= 2E[\text{Re}\{(y_0(n) - y_2(n))(y_2(n) - y_1(n))^*\} +$$
$$\lambda(n)\,|(y_2(n) - y_1(n))|^2]. \qquad (21)$$

Setting $\dfrac{\partial E[|e_a(n)|^2]}{\partial \lambda(n)} = 0$,

$$\lambda(n) = \frac{E[\text{Re}\{(d(n) - y_2(n))(y_2(n) - y_1(n))^*\}]}{E[|(y_2(n) - y_1(n))|^2]} \qquad (22)$$

where the unknown system output, $y_0(n)$, is replaced by $d(n)$, which is a valid assumption.

The mathematical expectation in the numerator and the denominator of (22) are replaced by exponential averaging

$$p_u(n) = (1 - \gamma)p_u(n-1) + \gamma u^2(n) \qquad (23)$$

where $p_u(n)$ is the averaged quantity, $u(n)$ is the signal to be averaged, and $\gamma = 0.01$. These results obtained for both
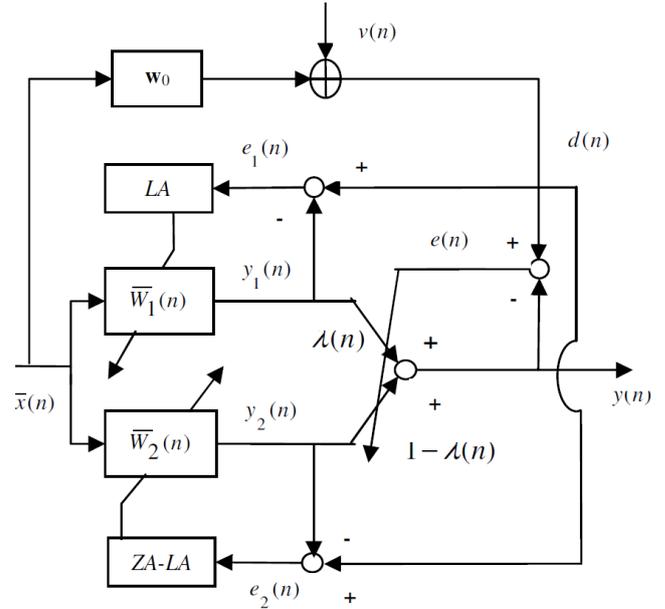


**Fig. 2.** Affine combination of two adaptive filters with mixing parameter $\lambda(n)$.

the numerator and denominator are substituted in (22) to obtain $\lambda(n)$.

# 6. Stability Analysis

Let us first provide the convergence analysis of LA filter. When the adaptive filter coefficient vector $\overline{W}(n)$ is updated by (3) and (5), the Lyapunov function is chosen as

$$V(n) = \beta^n e^2(n) . \qquad (24)$$

According to Lyapunov stability theory [30], the tracking error e(n) will asymptotically converge to zero.

$$e(n) = d(n) - W^T(n)\overline{x}(n) \qquad (25)$$
$$= d(n) - (\overline{W}^T(n-1) + g^T(n)\alpha(n))\overline{x}(n)$$
$$= d(n) - \overline{W}^T(n-1)\overline{x}(n) - g^T(n)\alpha(n)\overline{x}(n)$$
$$= \alpha(n) - g^T(n)\alpha(n)\overline{x}(n)$$
$$= \alpha(n) - \frac{\overline{x}^T(n)}{\|\overline{x}(n)\|^2}\left[1 - \frac{|e(n-1)|}{|\alpha(n)|\beta^{n/2}}\right]\alpha(n)\overline{x}(n)$$
$$= \alpha(n) - \alpha(n)\left[1 - \frac{|e(n-1)|}{|\alpha(n)|\beta^{n/2}}\right]$$
$$= \frac{\alpha(n)}{|\alpha(n)|}\frac{|e(n-1)|}{\beta^{n/2}}$$
$$= |e(n-1)|\beta^{-n/2}\text{sgn}(\alpha(n)) ,$$

$$\therefore |e(n)| = |e(n-1)|\beta^{-n/2} \qquad (26)$$
$$|e(1)| = |e(0)|\beta^{-1/2}$$
$$|e(2)| = |e(1)|\beta^{-2/2} = |e(0)|\beta^{-(2+1)/2}$$
$$\vdots$$
$$|e(n)| = |e(n-1)|\beta^{-n/2} = \cdots = |e(0)|\beta^{-(n+1)n/4} .$$

Therefore, the tracking error $e(n)$ converges to zero exponentially according to

$$|e(n)| = |e(0)| \beta^{-(n+1)n/4}. \tag{27}$$

Now, consider the output of the individual filters of ACLA filter which is expressed as

$$y_i(n) = \bar{W}_i^{\mathrm{T}}(n-1)\bar{x}(n), \quad i = 1, 2, \tag{28}$$

and the overall system error is given by

$$e(n) = d(n) - y(n) \tag{29}$$

where

$$d(n) = w_0^{\mathrm{T}} \bar{x}(n) + v(n) \tag{30}$$

and $y(n)$ is the ACLA filter output as given in (18).

Equation (18) can be written as

$$\begin{aligned} y(n) &= \lambda(n)\bar{W}_1^{\mathrm{T}}(n)\bar{x}(n) + [1-\lambda(n)]\bar{W}_2^{\mathrm{T}}(n)\bar{x}(n) \\ &= \left\{ \lambda(n)[\bar{W}_1(n) - \bar{W}_2(n)] + \bar{W}_2(n) \right\}^{\mathrm{T}} \bar{x}(n) \\ &= \left\{ \lambda(n)\bar{W}_{12}(n) + \bar{W}_2(n) \right\}^{\mathrm{T}} \bar{x}(n) \end{aligned} \tag{31}$$

where

$$\bar{W}_{12}(n) = \bar{W}_1(n) - \bar{W}_2(n). \tag{32}$$

The equivalent weight vector $\bar{W}_c$ of the combined filter can be expressed as

$$\begin{aligned} \bar{W}_c(n) &= \lambda(n)\bar{W}_1(n) + [1-\lambda(n)]\bar{W}_2(n) \\ &= \lambda(n)[\bar{W}_1(n) - W_2(n)] + \bar{W}_2(n) \\ &= \lambda(n)\bar{W}_{12}(n) + \bar{W}_2(n). \end{aligned} \tag{33}$$

Using (30) and (31) in (29), we get

$$\begin{aligned} e(n) &= w_0^{\mathrm{T}} \bar{x}(n) + v(n) - \left\{ \lambda(n)\bar{W}_{12}(n) + \bar{W}_2(n) \right\}^{\mathrm{T}} \bar{x}(n) \\ &= v(n) + \left[ \bar{W}_{02}(n) - \lambda(n)\bar{W}_{12}(n) \right]^{\mathrm{T}} \bar{x}(n) \end{aligned} \tag{34}$$

where

$$\bar{W}_{02}(n) = w_0 - \bar{W}_2(n). \tag{35}$$

The Mean Square Deviation (MSD) of the ACLA filter at time $n$ is

$$\begin{aligned} MSD_c(n) &= E\left\{ [w_0 - \bar{W}_c(n)]^{\mathrm{T}} [w_0 - \bar{W}_c(n)] \right\} \\ &= E\left\{ \left[ w_0 - \lambda(n)\bar{W}_{12}(n) - W_2(n) \right]^{\mathrm{T}} \left[ w_0 - \lambda(n)\bar{W}_{12}(n) - W_2(n) \right] \right\} \\ &= E\left\{ \left[ \bar{W}_{02}(n) - \lambda(n)\bar{W}_{12}(n) \right]^{\mathrm{T}} \left[ \bar{W}_{02}(n) - \lambda(n)\bar{W}_{12}(n) \right] \right\} \\ &= E\left[ \bar{W}_{02}^{\mathrm{T}}(n)\bar{W}_{02}(n) \right] - 2E\left[ \lambda(n)\bar{W}_{02}^{\mathrm{T}}(n)\bar{W}_{12}(n) \right] + \\ &\quad E\left[ \lambda^2(n)\bar{W}_{12}^{\mathrm{T}}(n)\bar{W}_{12}(n) \right]. \end{aligned} \tag{36}$$

Substituting (22) in (36) yields

$$MSD_c(n) = E\left[ \bar{W}_{02}^{\mathrm{T}}(n)\bar{W}_{02}(n) \right] - E\left\{ \frac{[\bar{W}_{02}^{\mathrm{T}}(n)\bar{W}_{12}(n)]^2}{\bar{W}_{12}^{\mathrm{T}}(n)\bar{W}_{12}(n)} \right\}. \tag{37}$$

The first term of (37) corresponds to the MSD of the second adaptive filter, $MSD_2(n)$ and since the MSD is a positive quantity, it indicates that $MSD_c(n)$ is smaller than $MSD_2(n)$.

Equation (33) can also be expressed as

$$\bar{W}_c(n) = \bar{W}_1(n) - [1-\lambda(n)]\bar{W}_{12}(n). \tag{38}$$

Now by inserting (38) in the first line of (36), it can be shown that

$$MSD_c(n) = E\left[ \bar{W}_{01}^{\mathrm{T}}(n)\bar{W}_{01}(n) \right] - E\left\{ \frac{[\bar{W}_{01}^{\mathrm{T}}(n)\bar{W}_{12}(n)]^2}{\bar{W}_{12}^{\mathrm{T}}(n)\bar{W}_{12}(n)} \right\} \tag{39}$$

where

$$\bar{W}_{01}(n) = w_0 - \bar{W}_1(n). \tag{40}$$

The first term of (39) corresponds to the MSD of the first adaptive filter, $MSD_1(n)$ and since the MSD is a positive quantity, it indicates that $MSD_c(n)$ is smaller than $MSD_1(n)$. Thus, the combined filter performs at least as well as the best component filter or better than any of them, for every $n$ and the stability is guaranteed.

# 7. Simulation Results

This section shows the simulations that are carried out to evaluate the performance of our proposed sparse algorithms. The length of unknown system $\mathbf{w}_0$ is set to $N = 16$, and its impulse response is assumed to have only one non-zero value at the tap index 5 and zeroes elsewhere making the system highly sparse as shown in Fig. 3.

The input signal $\bar{x}(n)$ is considered in two ways:

Case 1: Gaussian random signal with zero mean and unit variance, $N(0,1)$ and

Case 2: Correlated/Colored signal generated by passing a white Gaussian, $u(n)$ through a first-order autoregressive process, AR(1) with a pole 0.8 that is represented as $x(n) = 0.8\, x(n-1) + u(n)$.
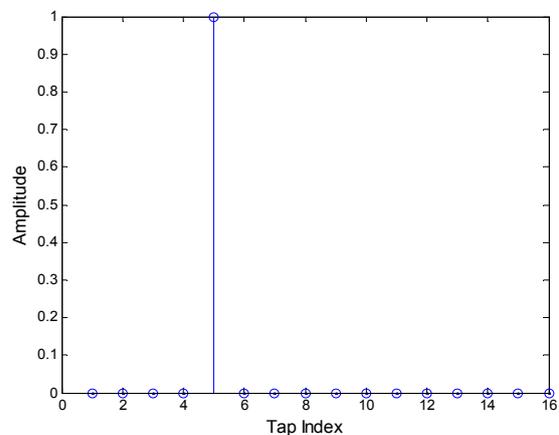


**Fig. 3.** Impulse response of a sparse system.

The output of the system is corrupted by an independent white Gaussian noise with variance 0.001.

The performance metrics used to evaluate the proposed algorithms are Mean Square Deviation (MSD), which is defined as

$$MSD\left\{\bar{W}(n)\right\} = E\left\{\left\|\bar{w}_0 - \bar{W}(n)\right\|_2^2\right\} \qquad (41)$$

where $E\{\bullet\}$ denotes expectation operator, and $\bar{w}_0$ and $\bar{W}(n)$ are the true FIR filter vector and its adaptive estimator, respectively, and the Mean Square Error (MSE) which is given as

$$MSE(n) = E\left\{e^2(n)\right\}. \qquad (42)$$

The average of 200 trials is used in evaluating the results.

From the simulation results shown in Fig. 4(a), for the Gaussian input and when the system is highly sparse, it is observed that our proposed Sparse LA filters (ZA-LA & RZA-LA) converge faster than the existing LMS and LA algorithms which cannot exploit the sparseness information present in the system. In Fig. 4(b), for the case of colored input, it is observed that the proposed LA algorithms have converged while the LMS algorithms fail to converge.

The mean square error (MSE) performance of the proposed LA algorithms is depicted in Fig. 5. It is observed that the MSE of proposed LA algorithms is lower than the LMS algorithm and their sparse counterpart.
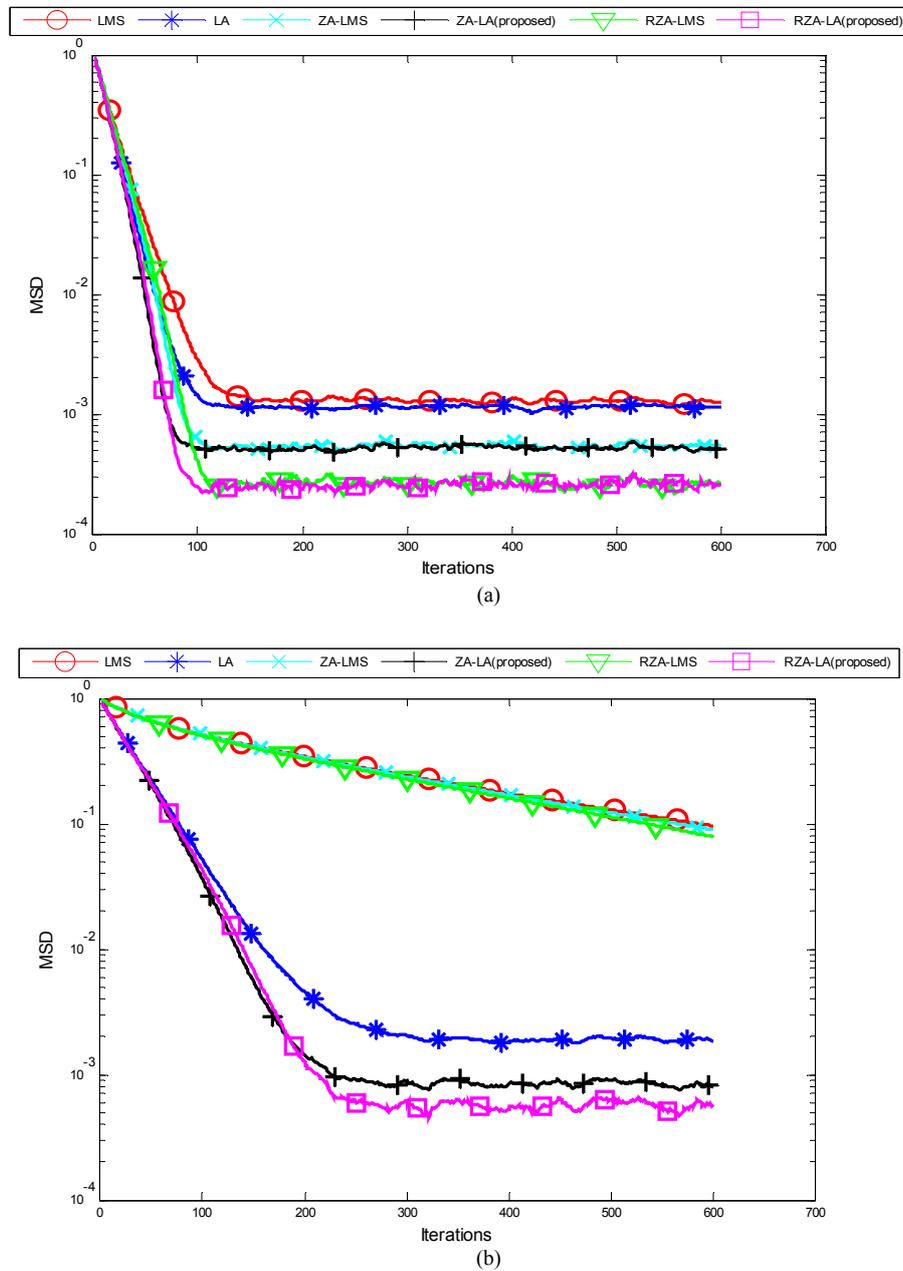


**Fig. 4.** MSD comparison of the proposed sparse LA algorithms with existing adaptive algorithms for highly sparse system with (a) white input, (b) colored input.
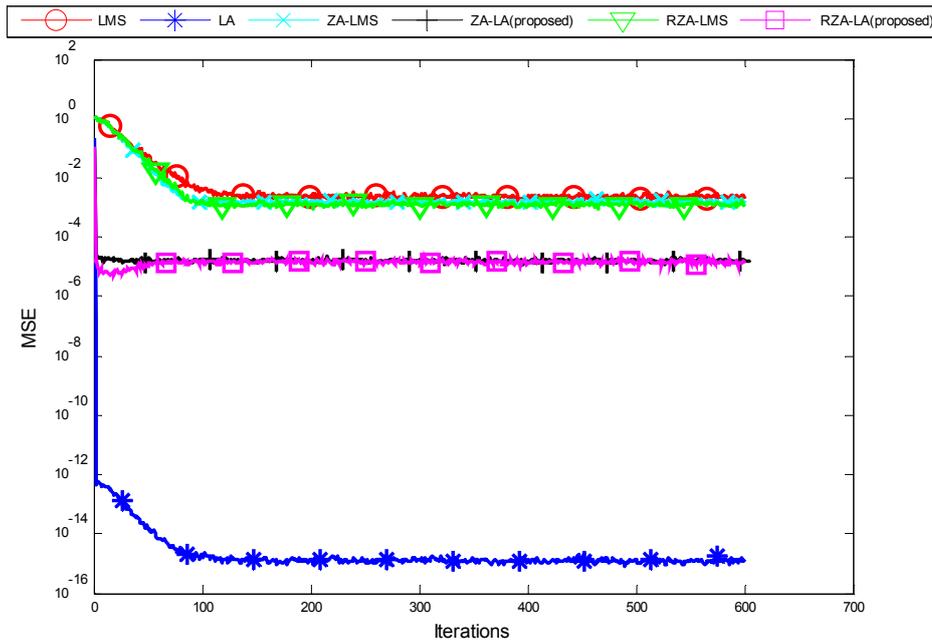
**Fig. 5.** MSE comparison of the proposed sparse LA algorithms with existing adaptive algorithms for highly sparse system.
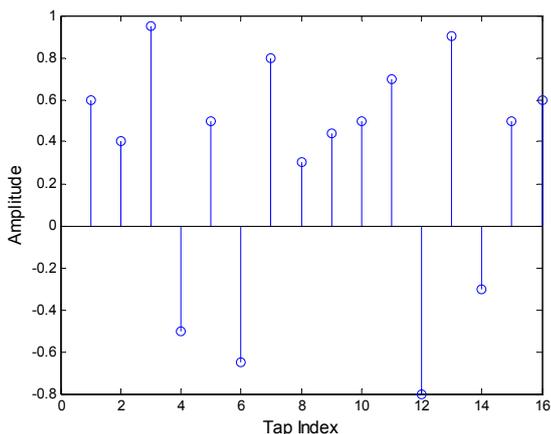


**Fig. 6.** Impulse response of a non-sparse system.

Next, we have considered that the unknown system has all its tap coefficients set to non-zero values, i.e. a non-sparse system. Figure 6 shows the impulse response of the used non-sparse system.

With the Gaussian input, the proposed ZA-LA algorithm converges faster than ZA-LMS algorithm, but it exhibits high steady state error when the system is non-sparse as shown in Fig. 7(a). The convergence of RZA-LA algorithm is the same as that of the LA algorithm and much better than the LMS algorithm. In the case of colored input, we observe from Fig. 7(b) that the performance of our proposed algorithms is superior to that of LMS algorithm and their sparse counterpart.

The mean square error (MSE) of proposed LA algorithms is shown in Fig. 8. The RZA-LA algorithm achieves lower MSE value than that of the ZA-LA algorithm which is lower than that of the LMS algorithm and its sparse counterpart.

The performance of the ACLA filter is analyzed for identifying the system of length $N = 16$ with variable sparsity. Initially, the system is assumed to be a highly sparse system with impulse response as shown in Fig. 3. At the $600^{th}$ sample, the system is abruptly changed to a non-sparse system with impulse response as shown in Fig. 6. The learning curves of ACLA filter are shown in Fig. 9 and Fig. 10 for white and colored input cases, respectively.

From Fig. 9, it can be seen clearly that for white input case, the ACLA filter achieves faster convergence and better steady state behavior. When the system is highly sparse (before the $600^{th}$ sample), the ACLA filter attains the lower steady state value of ZA-LA filter, and when the system is changed to a non-sparse system (at the $600^{th}$ sample), it achieves the steady state value of LA filter. When the input is considered as colored input, the steady state value of the ACLA filter is slightly better than that of the independent filters for a highly sparse system and when the system is changed to a non-sparse system, the ACLA filter behaves like LA filter that achieves the lower steady state value as shown in Fig. 10. Thus, the proposed ACLA filter is robust in identifying the systems with variable sparsity. Figure 11 shows the evolution of mixing parameter, $\lambda(n)$ for the ACLA filter.

The behavior of the proposed algorithms with input as a real speech signal sampled at 8 kHz is evaluated in the next simulation. The system is considered to have a varying degree of sparsity. Initially, the system is assumed to be sparse with impulse response shown in Fig. 3, and at time $t = 1.75$ sec, it is changed to the non-sparse system as shown in Fig. 6. The simulations are averaged over 100 trials. The Echo Return Loss Enhancement (ERLE) is used as the performance metric and is defined as [23], [37]

$$ERLE = 10 \log \frac{E[(d(n) - v(n)]^2}{E[(e(n) - v(n))]^2} \quad [\text{dB}]. \qquad (43)$$

It is observed from Fig. 12 that the proposed RZA-LA filter achieves higher ERLE values and the ERLE of ZA-LA filter is better than that of the LA filter when the sys-

tem is sparse. At $t = 1.75$ sec, when the system impulse response is switched to non-sparse, the ERLE performance of ZA-LA is reduced while the RZA-LA filter still attains the higher ERLE value closer to the LA filter which performs well under non-sparse conditions. Hence, the robustness of the proposed algorithms is verified.
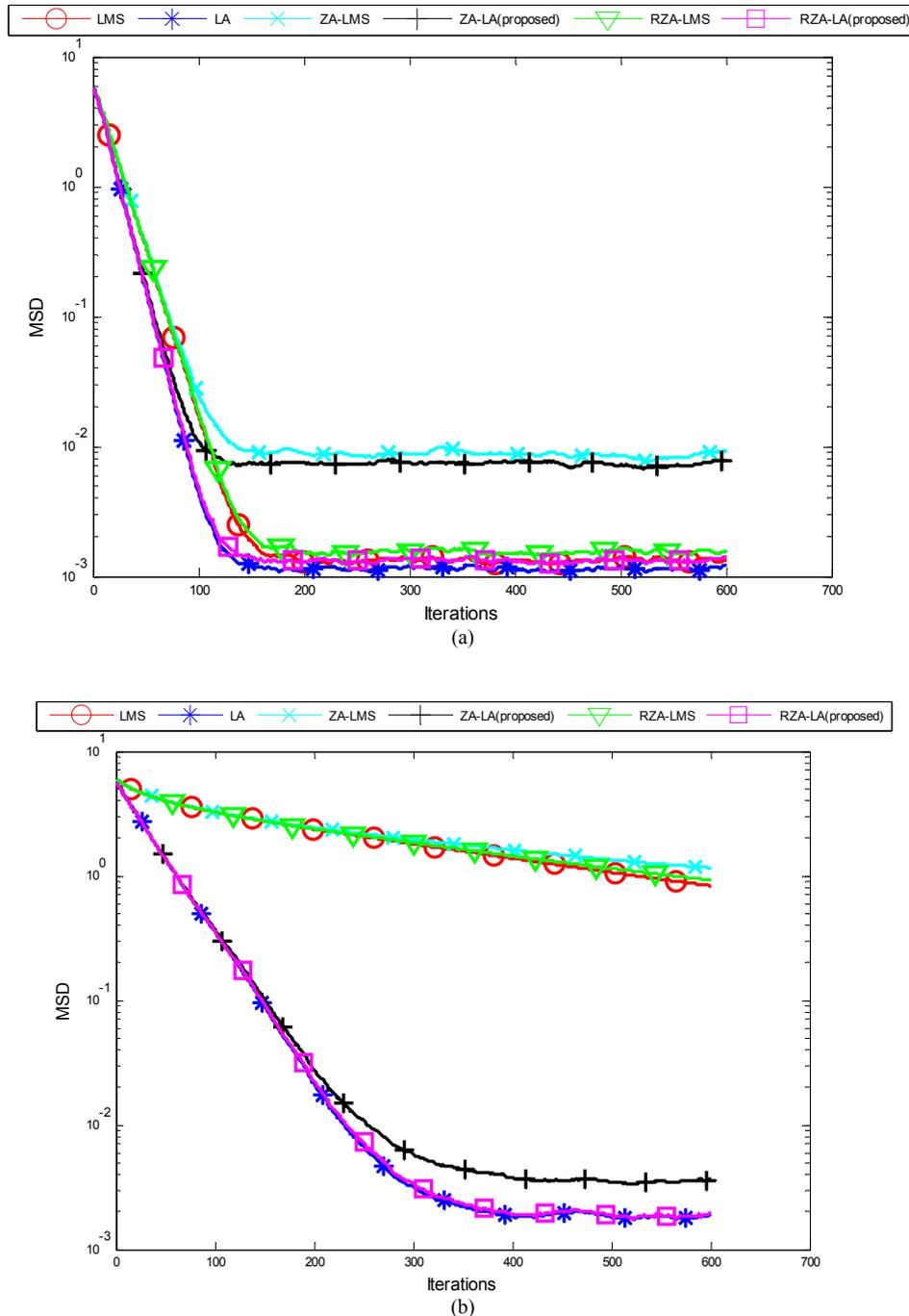


**Fig. 7.** MSD comparison of the proposed sparse LA algorithms with existing adaptive algorithms for non-sparse system with (a) white input, (b) colored input.
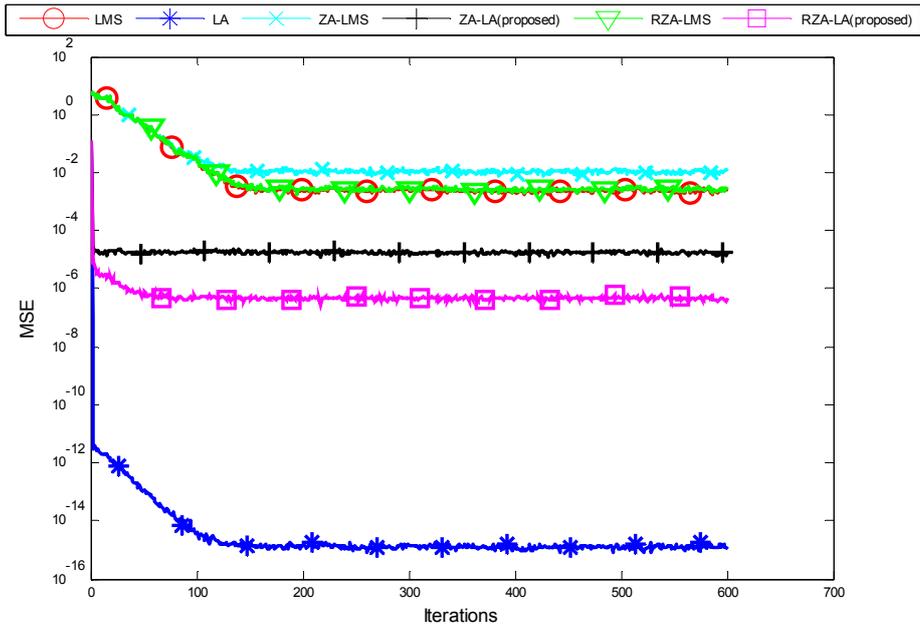
**Fig. 8.** MSE comparison of the proposed sparse LA algorithms with existing adaptive algorithms for non-sparse system.
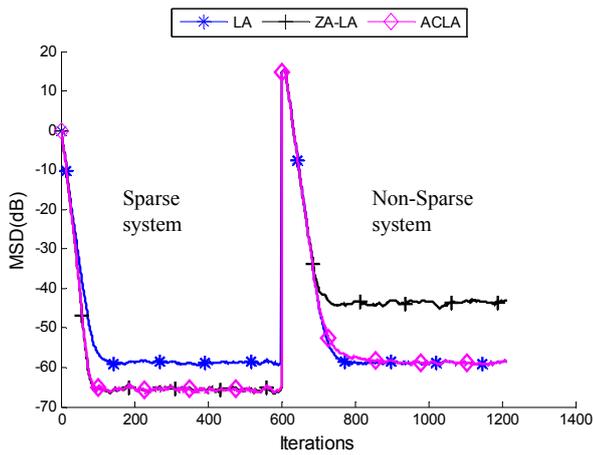


**Fig. 9.** Tracking and steady-state performance of ACLA filter for white input.
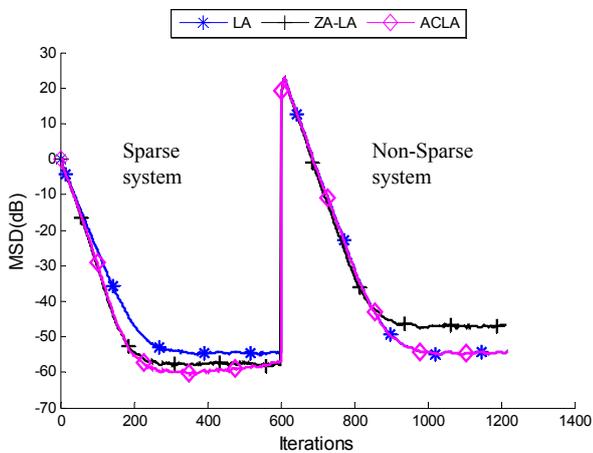


**Fig. 10.** Tracking and steady-state performance of ACLA filter for colored input.
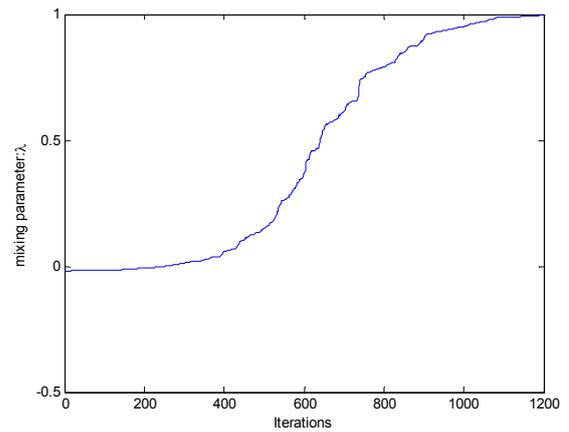


**Fig. 11.** Evolution of mixing parameter $\lambda(n)$ for the proposed affine combination.
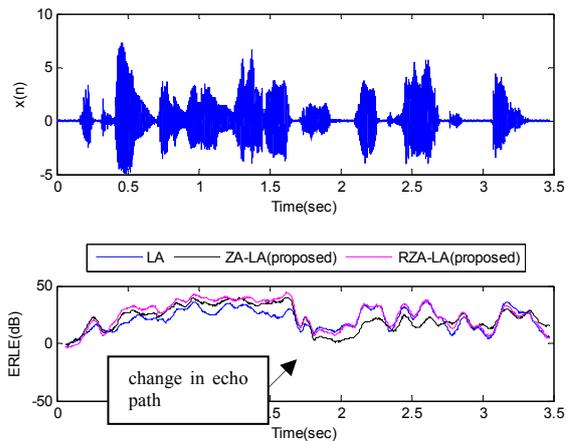


**Fig. 12.** ERLE performance plot. The input signal, x($n$) is a real speech signal sampled at 8 kHz (top). At $t = 1.75$ sec the system impulse response changes from sparse to non-sparse. The ERLE comparison of the proposed filters (bottom).

# 8. Conclusion

The standard Lyapunov Adaptation algorithm does not exploit sparsity present in an unknown sparse system. In this paper, we proposed two novel algorithms, namely, Zero-Attracting Lyapunov Adaptation algorithm (ZA-LA) and Reweighted Zero-Attracting Lyapunov Adaptation algorithm (RZA-LA) to improve adaptive sparse system identification performance. From the simulation results, the effectiveness of the proposed algorithms is verified for both white input and colored input case in terms of MSD and MSE.

Also, an Affine Combined Lyapunov Adaptation (ACLA) filter is presented to identify the systems with variable sparsity. The proposed combination filter exhibits robustness and achieves lower steady state value irrespective of the level of sparseness for the unknown system. The added complexity of the proposed algorithms is worth considering due to the increased performance.

# Acknowledgments

# References

[1] HAYKIN, S. *Adaptive Filter Theory*. 4th ed. Upper Saddle River (NJ, USA): Prentice Hall, 2002. ISBN: 9780130901262

[2] WIDROW, B. STEARNS, S. D. *Adaptive Signal Processing*. Englewood Cliffs (NJ, USA): Prentice Hall, 1985. ISBN: 0-13-004029-0

[3] MAN, Z., WU, H. R., LAI, W., et al. Design of adaptive filters using Lyapunov stability theory. In *Proceedings of the 6th IEEE International Workshop on Intelligent Signal Processing and Communication Systems*. Melbourne (Australia), 1998, p. 304–308. ISBN: 0732620309

[4] SENG, K. P., MAN, Z., WU, H. R. Lyapunov-theory-based radial basis function networks for adaptive filtering. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications,* 2002, vol. 49, no. 8, p. 1215–1220. DOI: 10.1109/TCSI.2002.801255

[5] MAN, Z., WU, H. R., LIU, S., et al. A new adaptive backpropagation algorithm based on Lyapunov stability theory for neural networks. *IEEE Transactions on Neural Networks,* 2006, vol. 17, no. 6, p. 1580–1581. DOI: 10.1109/TNN.2006.880360

[6] ZHAO, H., ZHANG, J. Filtered-s Lyapunov algorithm for active control of nonlinear noise processes. In *Proceedings of the 9th International Conference on Signal Processing*. Beijing (China), 2008, p. 311–314. DOI: 10.1109/ICOSP.2008.4697133

[7] RADECKI, J., ZILIC, Z., RADECKA, K. Echo cancellation in IP networks. In *Proceedings of the 45th Midwest Symposium on Circuits and Systems*. Tulsa (USA), August 2002, p. 219–222. DOI: 10.1109/MWSCAS.2002.1186837

[8] SCHREIBER, W. F. Advanced television systems for terrestrial broadcasting: some problems and some proposed solutions. In *Proceedings of the IEEE*, 1995, vol. 83, no. 6, p. 958–981. DOI: 10.1109/5.387095

[9] LI, W., PREISIG, J. C. Estimation of rapidly time-varying sparse channels. *IEEE Journal of Oceanic Engineering*, 2007, vol. 32, no. 4, p. 927–939. DOI: 10.1109/JOE.2007.906409

[10] SHIH, S. Y., CHEN. K. C. Compressed sensing construction of spectrum map for routing in cognitive radio networks. *Wireless Communication and Mobile Computing*, 2012, vol. 12, no. 18, p. 1592–1607. DOI: 10.1002/wcm.2338

[11] DAI, L., WANG, Z., YANG, Z. Compressive sensing based time domain synchronous OFDM transmission for vehicular communications. *IEEE Journal on Selected Areas in Communications*, 2013, vol. 31, no. 9, p. 460–469. DOI: 10.1109/JSAC.2013.SUP.0513041

[12] GUI, G., MEHBODNIYA, A., ADACHI, F. Bayesian sparse channel estimation and data detection for OFDM communication systems. In *Proceedings 2013 IEEE 78th Vehicular Technology Conference (VTC-Fall)*. LasVegas (USA), September 2013, p. 1–5. DOI: 10.1109/VTCFall.2013.6692113

[13] CANDES, E. Compressive sampling. In *Proceedings of the International Congress of Mathematics*. Madrid (Spain), 2006, vol. 3, p. 1433–1452. DOI: 10.4171/022-3/69

[14] BARANIUK, R. G. Compressive sensing. *IEEE Signal Processing Magazine*, July 2007, vol. 24, no. 4, p. 118–121. DOI: 10.1109/MSP.2007.4286571

[15] TIBSHIRANI, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society (B)*, 1996, vol. 58, no. 1, p. 267–288.

[16] CHEN, Y., GU, Y., HERO III, A. O. Sparse LMS for system identification. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei (Taiwan), 2009, p. 3125 to 3128. DOI: 10.1109/ICASSP.2009.4960286

[17] CANDES, E. J., WAKIN, M. B., BOYD, S. P. Enhancing sparsity by reweighted $\ell 1$ minimization. *Journal of Fourier Analysis and Applications*, 2008, vol. 14, no. 5, p. 877–905. DOI: 10.1007/s00041-008-9045-x

[18] MARTÍNEZ-RAMÓN, M., ARENAS-GARCÍA, J., NAVIA-VÁZQUEZ, A., et al. An adaptive combination of adaptive filters for plant identification. In *Proceedings of the 14th International Conference on DSP*. Santorini (Greece), July 2002, p. 1195–1198. DOI: 10.1109/ICDSP.2002.1028307

[19] ARENAS-GARCÍA, J., MARTÍNEZ-RAMÓN, M., GÓMEZ-VERDEJO, V., et al. Multiple plant identifier via adaptive LMS convex combination. In *Proceedings IEEE International Symposium on Intelligent Signal Processing*. Budapest (Hungary), 2003, p. 137–142. DOI: 10.1109/ISP.2003.1275828

[20] DAS, B. K., CHAKRABORTY, M. Sparse adaptive filtering by an adaptive convex combination of the LMS and the ZA-LMS algorithms. *IEEE Transactions on Circuits and Systems I*, 2014, vol. 61, no. 5, p. 1499–1507. DOI: 10.1109/TCSI.2013.2289407

[21] GOGINENI, V. C., CHAKRABORTY, M. Adaptive convex combination of APA and ZA-APA algorithms for sparse system identification. 19 pages. [Online] Cited 2015-09-10. Available at: https://arxiv.org/pdf/1509.03203.pdf

[22] RADHIKA, S., ARUMUGAM, S. Steady state mean square analysis of convex combination of ZA-APA and APA for acoustic echo cancellation. In *Proceedings of International Symposium on Intelligent Systems Technologies and Applications*. Kochi (India), 2015, p. 437–446. DOI: 10.1007/978-3-319-23036-8_38

[23] ARENAS-GARCÍA, J., FIGUEIRAS-VIDAL, A. R. Adaptive combination of proportionate filters for sparse echo cancellation.

*IEEE Transactions on Audio, Speech, and Language Processing*, August 2009, vol. 17, no. 6, p. 1087–1098. DOI: 10.1109/TASL.2009.2019925

[24] BERSHAD, N. J., BERMUDEZ, J. C. M., TOURNERET, J.-Y. An affine combination of two LMS adaptive filters—transient mean-square analysis. *IEEE Transactions on Signal Processing*, May 2008, vol. 56, no. 5, p. 1853–1864. DOI: 10.1109/TSP.2007.911486

[25] CANDIDO, R., SILVA, M. T. M., NASCIMENTO, V. H. Affine combinations of adaptive filters. In *42nd Asilomar Conference on Signals, Systems and Computers*. Pacific Grove (CA, USA), 2008. DOI: 10.1109/ACSSC.2008.5074399

[26] TRUMP, T. Output signal based combination of two NLMS adaptive filters – transient analysis. *Proceedings of Estonian Academy of Sciences*, 2011, vol. 60, no. 4, p. 258–268. DOI: 10.3176/proc.2011.4.06

[27] BUTSENKO, M., TRUMP, T. An affine combination of adaptive filters for sparse impulse response identification. In *23rd Telecommunications Forum (TELFOR)*. Belgrade, (Serbia), 2015, p. 396–399. DOI: 10.1109/TELFOR.2015.7377491

[28] BUTSENKO, M., TRUMP, T. An affine combination of adaptive filters for channels with different sparsity levels. *Telfor Journal*, 2016, vol. 8, no. 1, p. 32–37. DOI: 10.5937/telfor1601032B

[29] MULGREW, B., COWAN, C. F. N. *Adaptive Filters and Equalizers*. MA (USA): Kluwer Academic Publishers, 1988.

[30] SLOTINE, J. J. E., LI, W. *Applied Nonlinear Control*. NJ (USA): Prentice Hall, 1991. ISBN: 0-13-040890-5

[31] MENGÜC, E. C., ACIR, N. Lyapunov stability theory based adaptive filter algorithm for noisy measurements. In *2013 UKSim 15th International Conference on Computer Modelling and Simulation*. Cambridge (UK), 2013, p. 451–454. DOI: 10.1109/UKSim.2013.50

[32] MENGÜC, E. C., ACIR, N. A novel adaptive filter algorithm for tracking of chaotic time series. In *IEEE 19th Conference on Signal Processing and Communication Applications*. Antalya (Turkey), 2011, p. 490–493. (in Turkish), DOI: 10.1109/SIU.2011.5929694

[33] MENGÜC, E. C., ACIR, N. A novel adaptive filter design using Lyapunov stability theory. *Turkish Journal of Electrical Engineering & Computer Sciences*, 2015, vol. 23, no. 6, p. 719–728. DOI: 10.3906/elk-1212-29

[34] MENGÜC, E. C., ACIR. N. Real-time implementation of Lyapunov stability theory-based adaptive filter on FPGA. *IEICE Transactions on Electronics*, 2016, vol. E99.C, no. 1, p. 129–137. DOI: 10.1587/transele.E99.C.129

[35] FAZEL, M. Matrix rank minimization with applications. *Ph.D. Thesis*. Electrical Engineering Department, Stanford University, 2002, vol. 54, 130 p.

[36] FAZEL, M., HINDI, H., BOYD, S.P. Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices. In *Proceedings of the American Control Conference*, Denver, Colorado (USA). June 4-6, 2003, p. 2156–2162. DOI: 10.1109/ACC.2003.1243393

[37] HANS, M. C., LEVERGOOD, T. M. *Echo Cancelation*. Digital Equipment Corporation, Cambridge Research Lab, Technical report CRL 94/7, October 13, 1994.

## About the Authors ...

**Rakesh POGULA** received his B.Tech degree in Electronics & Communication Engineering from Jawaharlal Nehru Technological University, Hyderabad, India in 2006 and M.Tech in Communication Systems specialization from the National Institute of Technology Tiruchirappalli, Tamilnadu, India in 2009. Currently, he is a Ph.D. student in the Dept. of Electronics & Communication Engineering, National Institute of Technology Warangal, India. His research interests include adaptive signal processing, speech processing.

**T. Kishore KUMAR** received his Ph.D. in Digital Signal Processing from JNTU Hyderabad, India in 2004. He is currently working as an associate professor and head in the Dept. of E.C.E., National Institute of Technology Warangal, India. His research interests include speech signal processing, adaptive signal processing, etc.

**Felix ALBU** obtained the B.Sc. degree in Electronics and Ph.D. degree in Telecommunications from Politehnica University of Bucharest, Bucharest, Romania, in 1993 and 1999, respectively. He got the Dr. Habil. certificate in Electronics and Telecommunications in 2014. Currently, he is a Full Professor at the Valahia University of Targoviste, Romania. His research interests are speech/image processing, adaptive filtering, active noise control, neural networks, and machine learning.