Reconsider the Sparsity-Induced Least Mean Square Algorithms on Channel Estimation

Jie Wang(✉), Shangang Fan, Jie Yang, Jian Xiong, and Guan Gui

College of Telecommunication and Information Engineering,
Nanjing University of Posts and Telecommunications, Nanjing 210003, China
13675171572@163.com, sponder@126.com,
{jyang, jxiong, guiguan}@niupt.edu.cn

Abstract. This paper surveys recent advances related to sparse least mean square (LMS) algorithms. Since standard LMS algorithm does not take advantage of the sparsity information about the channel being estimated, various sparse LMS algorithms that are aim at outperforming standard LMS in sparse channel estimation are discussed. Sparse LMS algorithms force the solution to be sparse by introducing a sparse penalty to the standard LMS cost function. Under the reasonable conditions on the training datas and parameters, sparse LMS algorithms are shown to be mean square stable, and their mean square error performance and convergence rate are better than standard LMS algorithm. We introduce the sparse algorithms under Gaussian noises model. The simulation results presented in this work are useful in comparing sparse LMS algorithms against each other, and in comparing sparse LMS algorithms against standard LMS algorithm.

Keywords: Gradient decent · Least mean square (LMS) algorithm Sparse penalty · Sparse channel estimation · Gaussian noises model

1 Introduction

We all know that least mean square (LMS) algorithm has been applied to solve various technical problems in signal processing fields and broadband wireless communication areas including adaptive communication line enhancement [1], system identification [2], channel estimation [3], echo cancelation [4], etc. The main reasons are that LMS algorithm has low computational complexity as well as does not need extensive stochastic knowledges of the channel models and the training data sequences compared to some other parameter estimation methods, for example, the recursive least squares (RLS) algorithm [5]. But the standard LMS algorithm don't consider the inherent sparse structure information of the system model which must weaken the performance of the estimation. The sparse channel means that there are a few domain taps, in another words, more than half of the channel coefficients are zero or near to zero [6–10]. The sparse channel structure is shown as in Fig. 1 in which the length of the channel is 16 but only four taps are nonzero. To utilize the strengths of LMS algorithm and take full advantage of the sparsity of the channel, many sparse LMS algorithms have been proposed in recent years. This paper main analyses the performance of spare channel estimation based on the sparse LMS algorithms under additive white Gaussian noises model.

C. Li and S. Mao (Eds.): WiCON 2017, LNICST 230, pp. 85–102, 2018.
https://doi.org/10.1007/978-3-319-90802-1_8

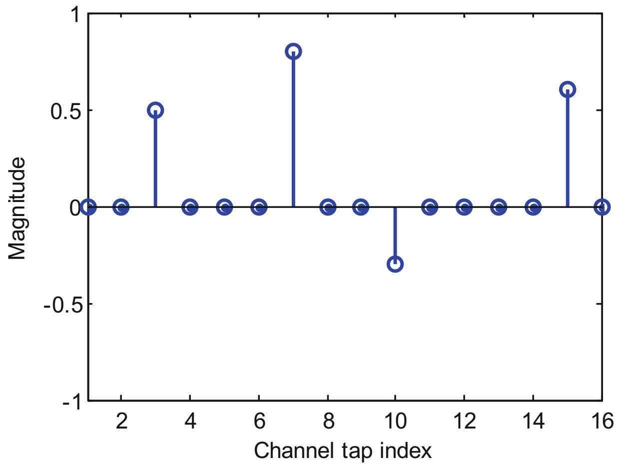

Fig. 1. Example of typical sparse channel model.

The proposed sparse LMS algorithms force the estimation to be sparse by introducing a sparse penalty to the classical LMS cost function. The sparse constraint forces the small channel coefficients to zero, which speed up convergence rate and lower the steady state error of the estimation when most taps of the channel are zero. This paper surveys the field of sparse channel estimation based on sparse LMS algorithms and how to design the sparse penalty to achieve the better performance including faster convergence rate, smaller mean square error, lower computation complexity etc.

Zero attracting least mean square (ZA-LMS) algorithm and Reweighted zero attracting least mean square (RZA-LMS) algorithm have been proposed in [11]. Reweighted ℓ_1-norm penalized least mean square (RL1-LMS) algorithm and ℓ_p-norm penalized least mean square (ℓ_p-norm LMS) algorithm in [12] are the improvements of ZA-LMS. To make much better use of the sparse structure, ℓ_0-norm penalized least mean square (ℓ_0-norm LMS) algorithm has been proposed in [13].

In order to compare these sparse LMS algorithms against each other better, we organize the presentation of the article into four main components. Section 2 reviews the system model being estimated and the standard LMS algorithm. Performance analysis of different sparsity-aware LMS algorithms under Gaussian noises model are presented in Sect. 3. Computer simulation results are shown in Sect. 4. Section 5 concludes the paper.

2 System Model and Problem Formulation

2.1 System Model

The system model of the sparse channel being estimated is as shown in Fig. 2. $\mathbf{x}_n$ is the input signal sequence which is defined as $\mathbf{x}_n = [x(n), x(n-1), \ldots, x(n-N-1)]^T$. N denotes the length of the channel. $\mathbf{w}$ is the actual unknown channel vector which is defined as $\mathbf{w} = [w_0, w_1, \ldots, w_{N-1}]^{\mathrm{T}}$. $y(n)$ is the output signal of the actual channel, which is defined as $y(n) = \mathbf{w}^T\mathbf{x}_n$. $v(n)$ is the additive noise in the sparse channel being estimated. So $d(n)$ is the desired signal at the receiver side which is defined as

$$d(n) = \mathbf{w}^T\mathbf{x}_n + v(n) \tag{1}$$

And $\widehat{\mathbf{w}}_n$ presents the estimated channel vector at iteration n, which is defined as $\widehat{\mathbf{w}}_n = [w_{0,n}, w_{1,n}, \ldots, w_{N-1,n}]$. $\widehat{y}(n)$ is the output signal under the estimated channel vector $\widehat{\mathbf{w}}_n$, which is defined as $\widehat{y}(n) = \widehat{\mathbf{w}}_n^T\mathbf{x}_n$. $e(n)$ is the error between the desired output signal $d(n)$ and the output signal $\widehat{y}(n)$ based on the estimated channel coefficients, which is given by

$$e(n) = d(n) - \widehat{y}(n) = \mathbf{w}^T\mathbf{x}_n - \widehat{\mathbf{w}}_n^T\mathbf{x}_n + v(n) \tag{2}$$

What we should do is to find a vector, which is given by $\widehat{\mathbf{w}}_n$ to make $e(n)$ minimum.

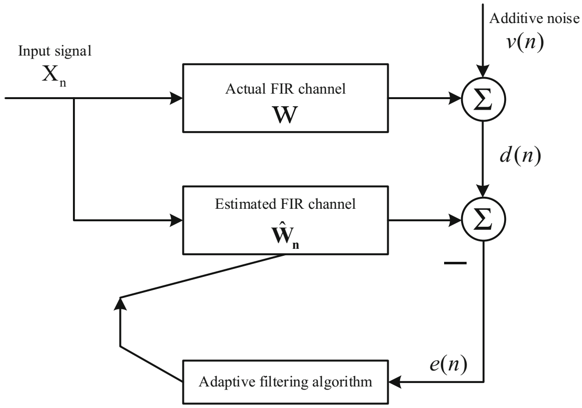

Fig. 2. LMS based system model.

2.2 Standard Least Mean Square Algorithm

Different adaptive filter algorithms utilize various cost functions to solve the problems of signal processing. We all seek the better cost functions either to adapt the channel better or to achieve the faster convergence speed. The cost function of standard least mean square algorithm is

$$G(n) = \frac{1}{2}e(n)^2 \tag{3}$$

And it use the gradient descent algorithm to minimize the Eq. (3) to get the solution of the actual channel vector [5]. So the iterative equation of the standard least mean square algorithm is given by

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \mu \nabla_{\mathbf{w}_n} G(n) = \mathbf{w}_n + \mu e(n) \mathbf{x}_n \tag{4}$$

where μ is the step size of the adaptive filter algorithm. The parameter μ is critical to guarantee the convergence and stable state of the adaptive algorithm. The lager μ, the faster convergence speed of the algorithm but the larger mean square divation (MSD), i.e. $\mathrm{MSD(n)} = \mathrm{E}\left\{\|\mathbf{w} - \widehat{\mathbf{w}}_n\|_2^2\right\}$. And vice versa. To make sure that the least mean square algorithm convergences, μ is chosen within the scope of $0 < \mu < \lambda_{\max}^{-1}$ with $\lambda_{\max}$ being the maximum eigenvalue of the covariance matrix of $\mathbf{x_n}$, i.e. $\mathbf{R} \triangleq \mathrm{E}\left[\mathbf{x}_n \mathbf{x}_n^{\mathbf{T}}\right]$ [5]. To see the influence of parameter μ directly, the simulation results with different μ is shown in Fig. 3. In the numerical simulation, we set the length of the channel being estimated N = 128, the input signal power to additive noise power ratio $SNR = 20\,\mathrm{dB}$. The weights vector of the system is random sequence. The input signal sequences are pseudo random binary sequences and $v(n)$ is additive white Gaussian noises. Three different μ values of 0.01, 0.008 and 0.005 are considered. Simulation results are obtained by taking the average of the network mean square error (MSE) over 2000 independent Monte Carlo runs to smooth the out curves. From Fig. 3 we can see that when μ is smaller the convergence rate is slower but the MSE is smaller and vice versa.

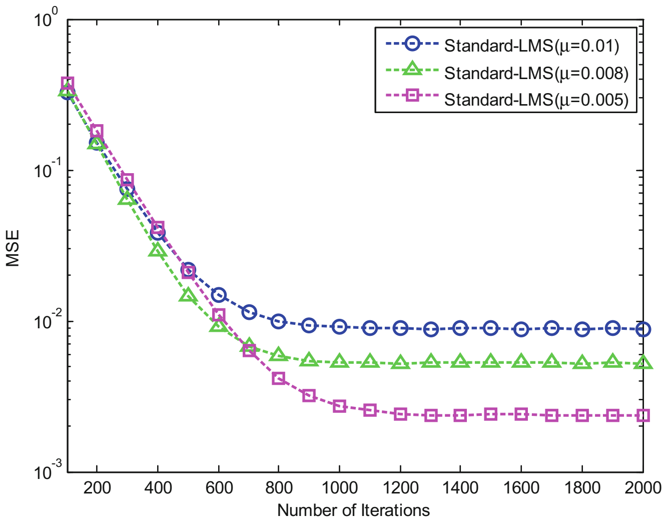

Fig. 3. MSE of standard LMS algorithm with different $\boldsymbol{\mu}$ values (SNR = 20 dB).

3 Reconsider of Sparse LMS Algorithm

In this section, all analysis are under the assumption that the additive noise $v(n)$ is additive white Gaussian noises model which is independent with input signal $\mathbf{x}_n$ [14].

3.1 The Presentation of Different Sparse Mean Square Algorithms

We present the system model and standard mean square algorithm in Section 2. In this Section, we analysis and compare the performance of different sparse adaptive filter algorithms under Gaussian noise environments. The basic idea of all the sparse mean square algorithms is to introduce a sparse penalty to the cost function of the standard LMS algorithm. The sparse constraint attracts the entries of the weights vector of the channel to zero in varying degrees. Accordingly, the estimation of the channel will have faster convergence rate and lower steady state error because most taps of the sparse channel are zero. The essential difference between various sparse algorithms is that the sparse penalties being proposed are diverse which will be analyzed in detail in the part B.

Zero attracting least mean square (ZA-LMS) algorithm: The cost function of ZA-LMS algorithm is $\mathrm{G}_{ZA}(\mathrm{n}) = \frac{1}{2}e^2(n) + \gamma_{ZA}\|\mathbf{w}_n\|_1$. Where $\gamma_{ZA}\|\mathbf{w}_n\|_1$ is sparse penalty, in which γ_{ZA} is regular parameter that balance the mean square error of the algorithm and the sparse degree of the system model. $\|.\|_1$ stands for the ℓ_1-norm of the vector. Based on the gradient decent algorithm, the update equation of the ZA-LMS algorithm is

$$\mathbf{w}_{\boldsymbol{n}+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}_n - \rho_{ZA} sgn(\mathbf{w}_n) \tag{5}$$

where $\rho_{\mathrm{ZA}} = \mu\gamma_{ZA}$.

Reweighted zero attracting least mean square (RZA-LMS) algorithm: The cost function of RZA-LMS is $\mathrm{G}_{\mathrm{RZA}}(n) = \frac{1}{2}e^2(n) + \gamma_{RZA}\sum_{i=1}^{N}\log\left(1 + [\mathbf{w}_n]_i/\epsilon'_{RZA}\right)$. Where $\gamma_{RZA}\sum_{i=1}^{N}\log\left(1 + [\mathbf{w}_n]_i/\epsilon'_{RZA}\right)$ is the sparse constraint, in which γ_{RZA} is regularization parameter that weights the mean square error of the algorithm and the sparse level of the system model and ϵ'_{RZA} is a positive number. Based on the gradient decent algorithm, the update equation of the RZA-LMS algorithm is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}_n - \rho_{RZA}\frac{sgn(\mathbf{w}_n)}{1 + \epsilon_{RZA}|\mathbf{w}_n|} \tag{6}$$

where $\rho_{RZA} = \mu\gamma_{RZA}\epsilon_{RZA}$ and $\epsilon_{\mathrm{RZA}} = 1/\epsilon'_{RZA}$.

Reweighted ℓ_1-norm penalized least mean square (RL1-LMS) algorithm: The cost function of RL1-LMS is $\mathrm{G}_{r\ell_1}(n) = \frac{1}{2}e^2(n) + \gamma_r\|\mathbf{s}_k\mathbf{w}_n\|_1$. Where $\gamma_r\|\mathbf{s}_k\mathbf{w}_n\|_1$ is the sparse constraint, in which γ_r is regular parameter that balance the mean square error of the algorithm and the sparse degree of the system model and $\mathbf{s}_k$ is a row vector with its elements are $[\mathbf{s}_k]_{\mathrm{i}} = \frac{1}{\epsilon_r + |[\mathbf{w}_{k-1}]_i|}$, i = 1, …, N. And ϵ_r is a positive number and $[.]_{\mathrm{i}}$ denotes the i-th entry of the vector. $\|.\|_1$ stands for the ℓ_1-norm of the vector. Based on the gradient decent algorithm, the update equation of the RL1-LMS is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}_n - \rho_r \frac{sgn(\mathbf{w}_n)}{\epsilon_r + |\mathbf{w}_{n-1}|} \tag{7}$$

where $\rho_r = \mu\gamma_r$ and sgn(.) is the sign function which can be presented by $\mathrm{sgn(x)} = \begin{cases} 1, x > 0 \\ 0, x = 0 \\ -1, x < 0 \end{cases}$. The operation of sgn(.) is on every entry of the vector. The absolute value operator and the division operator in the last term of Eq. (7) are all component-wise.

ℓ_p-norm penalized least mean square (ℓ_p-norm LMS) algorithm: The cost function of ℓ_p-norm LMS is $\mathrm{G}_{\ell_\mathrm{p}}(n) = \frac{1}{2}e^2(n) + \gamma_p\|\mathbf{w}_n\|_\mathrm{p}$. Where $\gamma_\mathrm{p}\|\mathbf{w}_n\|_\mathrm{p}$ is sparse penalty, in which γ_p is regular parameter that weights the mean square error of the algorithm and the sparse degree of the system model. $\|.\|_\mathrm{p}$ stands for the ℓ_p-norm of the vector. The parameter p is a positive number with $0 < p < 1$. Based on the gradient decent algorithm, the update equation of the ℓ_p-norm LMS algorithm is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}_n - \rho_p \frac{\left(\|\mathbf{w}_n\|_\mathrm{p}\right)^{1-\mathrm{p}} sgn(\mathbf{w}_n)}{|\mathbf{w}_n|^{(1-\mathrm{p})}} \tag{8}$$

But to prevent the algorithm from being unstable when $\mathbf{w}_n$ is a zero vector, we usually add a regularization parameter ϵ_p to the last term of Eq. (8). Then the update equation of the ℓ_p-norm LMS is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}_n - \rho_p \frac{\left(\|\mathbf{w}_n\|_p\right)^{1-p} sgn(\mathbf{w}_n)}{\epsilon_p + |\mathbf{w}_n|^{(1-p)}} \tag{9}$$

where $\rho_p = \mu\gamma_p$.

ℓ_0-norm penalized least mean square (ℓ_0-norm LMS) algorithm: The risk function of ℓ_0-norm LMS is $\mathrm{G}_{\ell_0}(n) = \frac{1}{2}e^2(n) + \gamma_{\ell_0}\|\mathbf{w}_n\|_0$. Where $\gamma_{\ell_0}\|\mathbf{w}_n\|_0$ is sparse penalty, in which γ_{ℓ_0} is regular parameter that balance the mean square error of the algorithm and the sparse degree of the system model. $\|.\|_0$ stands for the ℓ_0-norm of the vector. As we all know, find the minimum solution of the ℓ_0-norm is a Non-Polynomial (NP) hard problem. So a approximate continuous function has been proposed in [15] which is $\|\mathbf{w}\|_0 \approx \sum_{i=1}^{N}\left(1 - e^{-\beta\left|[\mathbf{w}_n]_\mathrm{i}\right|}\right)$. Where $|.|$ stands for the absolute operator and $[\mathbf{w}_n]_\mathrm{i}$ is the i-th element of the vector $\mathbf{w}$. β is a positive number. Then the cost function of ℓ_0-norm LMS can be rewritten as $\mathrm{G}_{\ell_0}(n) = \frac{1}{2}e^2(n) + \gamma_{\ell_0}\sum_{i=1}^{N}\left(1 - e^{-\beta\left|[\mathbf{w}_n]_\mathrm{i}\right|}\right)$. Based on the gradient decent algorithm, the update equation of the ℓ_0-norm LMS algorithm is

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}_n - \rho_{l_0}\beta\,\mathrm{sgn}(\mathbf{w}_n)\mathrm{e}^{-\beta|\mathbf{w}_n|} \tag{10}$$

where $\rho_{l_0} = \mu\gamma_{\ell_0}$. But the term $\mathrm{e}^{-\beta|\mathbf{w}_n|}$ has very high computational complexity which isn't what we want. Then the first-order Taylor series expansion of exponential functions has been introduced in [15] which is shown as $e^{-\beta|[\mathbf{w}_n]_i|} \approx f(x) = \begin{cases} 1-\beta|[\mathbf{w}_n]_i|, & when|[\mathbf{w}_n]_i| \le 1/\beta \\ 0, & others \end{cases}$. Then the update equation of the ℓ_0-norm LMS is derived as

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu e(n)\mathbf{x}_n - \rho_{l_0}\mathrm{J}(\mathbf{w}_n) \tag{11}$$

where $J([\mathbf{w}_n]_i) \approx \begin{cases} \beta sgn([\mathbf{w}_n]_i) - \beta^2([\mathbf{w}_n]_i), & when|[\mathbf{w}_n]_i| \le 1/\beta \\ 0, & others \end{cases}$, in which $[\mathbf{w}_n]_i$ is the i-th entry of the vector $\mathbf{w}_n$.

3.2 Analysis of the Difference Between the Five Sparse Algorithms

We present five sparse least mean square algorithms in part A. They have different sparse penalties which stand for the different sparse constraints to the channel vector being estimated. The regularization parameters in every algorithm also play a big role in the sparse constraint to the resolution of the channel.

The sparse penalty term of ZA-LMS algorithm is $\mathrm{h(t)}^{\mathrm{ZA}} = |\rho_{ZA} sgn(t)|$, where ρ_{ZA} decides the strength of the penalty term. From Fig. 4, we can know that ZA-LMS has the same sparse effects on all weights of the channel. It will cause the high steady state error of the algorithm against the dominant taps of the channel. As the improvement of ZA-LMS algorithm, RZA-LMS algorithm has a different sparse constraint which is $\mathrm{h(t)} = \left|\rho_{RZA}\frac{sgn(t)}{1+\epsilon_{RZA}|t|}\right|$, where ρ_{RZA} has the same role as the ρ_{ZA} plays. From Fig. 5, we can see that RZA-LMS algorithm attracts the entries of the channel vector to zero with different extent. The smaller the entry, the stronger strength. And vice versa. Accordingly, RZA-LMS algorithm normally have the better performance than ZA-LMS algorithm. The sparse penalty terms of the RL1-LMS algorithm and ℓ_p-norm LMS algorithm run similar as RZA-LMS algorithm. The main difference between the three algorithms is the distribution of the strength of sparse constraint to various taps of the channel. From Figs. 6 and 7, we can find that RL1-LMS algorithm and ℓ_p-norm LMS have a significantly stronger effects on small weights of the channel than RZA-LMS that will yield lower steady state error for sparse channel estimation.

The sparse penalty term of ℓ_0-norm LMS is $\mathrm{h(t)}^{\ell^0} = |\rho_{\ell_0}J(t)|$, where $J(t) \approx \begin{cases} \beta sgn(t) - \beta^2(t), & when|\mathrm{t}| \le 1/\beta \\ 0, & others \end{cases}$. From Fig. 8, we can get that the sparse attraction of ℓ_0-norm LMS algorithm only effects on the taps of the channel in a definite interval. This property of ℓ_0-norm LMS algorithm which reduce the mean square error of the weights out of the interval usually make it gain a more accurate estimation than another four sparse algorithms.

To compare the performance of the five sparse algorithms, we do the numerical simulation. The simulation results are shown in Fig. 9. The parameter values of every algorithm in the numeric simulation are shown in the Table 1.

From Fig. 9, we can see that the ZA-LMS have the faster convergence rate and larger MSE compared with the other four algorithm. And ℓ_0-norm LMS achieve the best performance in the simulation.

Table 1. Parameter values setting in Fig. 9.

Parameters	Values
Input signal	Pseudo-random binary sequences
The length of channel	N = 128
Number of dominant taps of channel vector	S = 16
Distribution of nonzero coefficients	Random Gaussian CN(0,1)
Signal to noise ratio for channel model	SNR = 20 dB
Gaussian noise distribution	Gaussian $\mathrm{CN}\left(0, \delta_n^2\right)$
Step size	$\mu = 0.005$
Parameters of algorithms	ρ_{ZA} = 1e−3, ρ_{RZA} = 1e−5, E_{RZA} = 25, ρ_r = 5e−6, p = 0.5, ρ_p = 5e−6, E_p = 0.05, β_0 = 10, ρ_0 = 1e−4

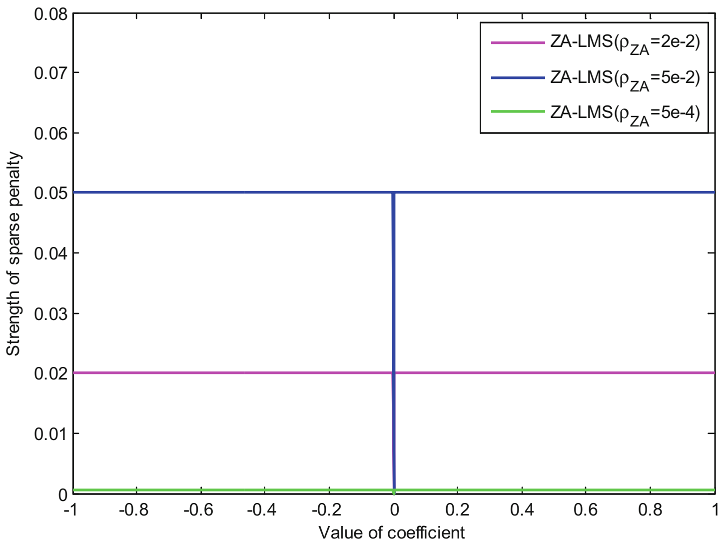

Fig. 4. Sparse constraints to channel vector **w** of ZA-LMS algorithm with different ρ_{ZA} values.

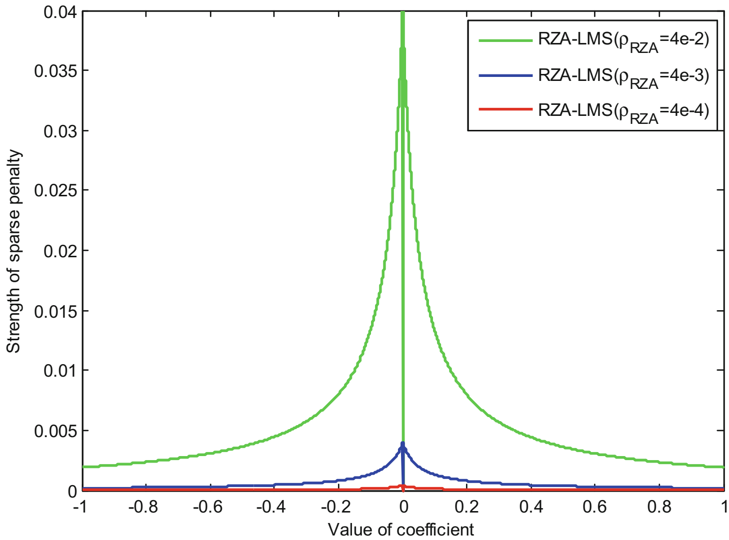

Fig. 5. Sparse constraints to channel vector **w** of RZA-LMS algorithm with different $\boldsymbol{\rho_{RZA}}$ values ($\boldsymbol{\epsilon_{RZA}} = \mathbf{20}$).

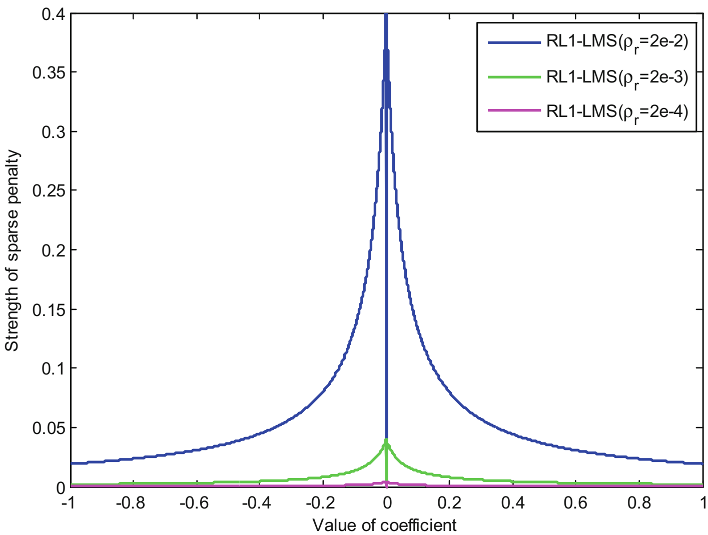

Fig. 6. Sparse constraints to channel vector **w** of RL1-LMS algorithm with different $\boldsymbol{\rho_r}$ values ($\boldsymbol{\epsilon_r} = \mathbf{5 \times 10^{-2}}$).

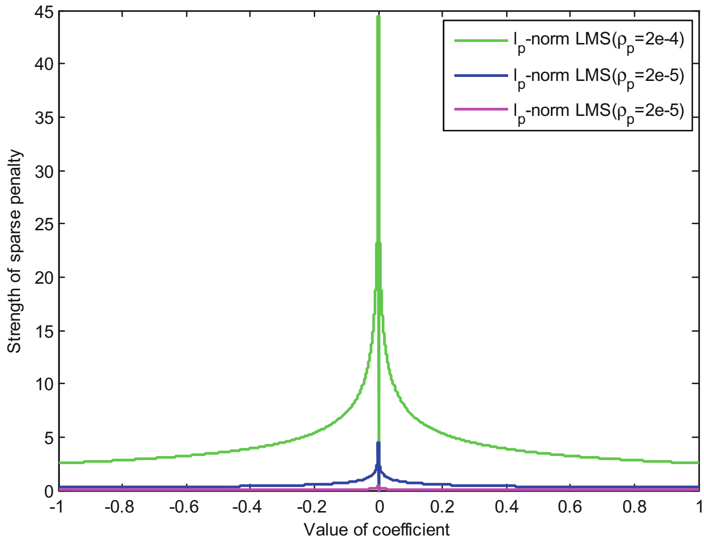

Fig. 7. Sparse constraints to channel vector **w** of ℓ_p-norm LMS algorithm with different $\boldsymbol{\rho_p}$ values ($\boldsymbol{\epsilon_p = 5 \times 10^{-2}}$,$\mathbf{p = 0.5}$).

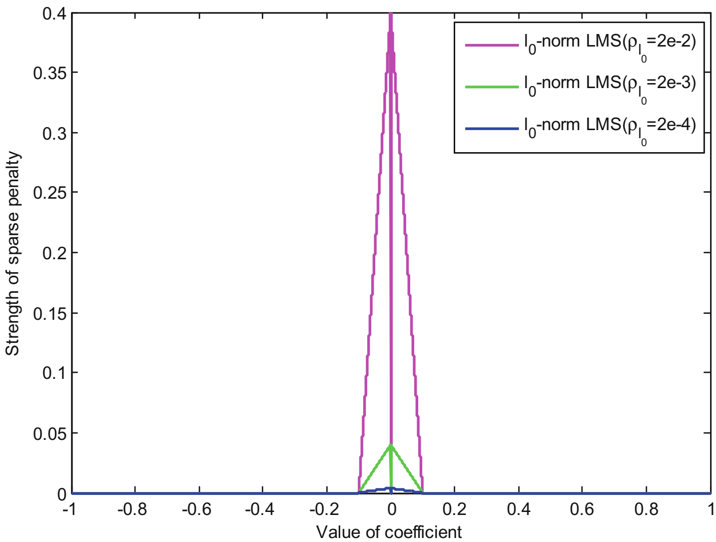

Fig. 8. Sparse constraints to channel vector **w** of ℓ_0-norm LMS algorithm with different $\boldsymbol{\rho_r}$ values ($\boldsymbol{\beta = 10}$).

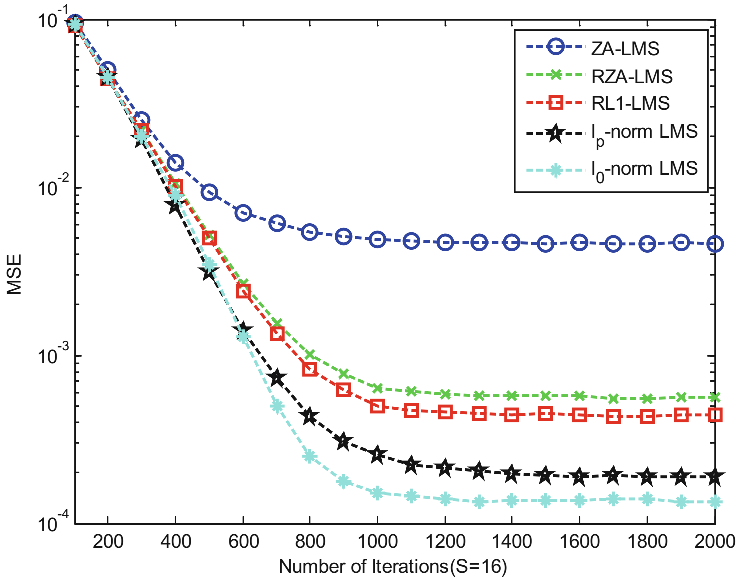

Fig. 9. Comparisons between different sparse LMS algorithms.

4 Simulation Results

Regular parameters of sparse LMS algorithms and the sparse degree of the channel have big implications for the gains of the algorithms. The steady state MSD equation $J_{ZA}(\infty)$ of ZA-LMS which has been derived in [16] is that

$$J_{ZA}(\infty) = \left[t^2 - \frac{(\pi-1)\mu\sigma_x^2 + 1}{2\pi\mu^2\sigma_x^4}\rho_{ZA}^2\right]\frac{2}{\mu\sigma_x^2} - \frac{\sigma_v^2}{\sigma_x^2} \tag{12}$$

where σ_v^2 denotes the variance of the additive noise $v(n)$ and σ_x^2 is the variance of the input sequence $\mathbf{x}(n)$. The parameter t is the positive solution of $c_1 t^2 + c_2 t + c_3 = 0$ in which $c_1 = L - \frac{2}{\mu\sigma_x^2}\left(1 - \mu\sigma_x^2\right)$ and $c_2 = -2(L-S)\rho_{ZA}\frac{\sqrt{1-\mu\sigma_x^2}}{\sqrt{2\pi\mu\sigma_x^2}}$ and $c_3 = \left(\frac{L-2S}{2\pi} + M + 1\right)\frac{1-\mu\sigma_x^2}{\mu^2\sigma_x^4}\rho_{ZA}^2 + \frac{\left(1-\mu\sigma_x^2\right)^2}{\pi\mu^3\sigma_x^6}\rho_{ZA}^2 + \frac{\sigma_v^2}{\sigma_x^2}\left(1 - \mu\sigma_x^2\right)$. From Eq. (12) we can see that the steady state MSD depends on the sparse degree of the channel S, the regular parameter of the algorithm ρ_{ZA}, the step size μ and the channel length L [16]. When the sparsity of the channel is determined, the larger value of ρ_{ZA} will increase the gap between the large tap-weights and it's true value and the smaller value of ρ_{ZA} will decrease the effects of sparse constraint on sparse channels. Both of them destroy the performance of ZA-LMS algorithm. To see the influence of ρ_{ZA} directly we simulate the algorithm with different ρ_{ZA} values. The simulation based on the channel model with the length of the channel is N = 128, the sparse level of the channel is S = 16, the signal to noise ratio of the channel is SNR = 20 dB. Simulation results are obtained by taking the average of the network mean square error (MSE) over 2000 independent Monte Carlo runs. The simulation results are shown in Fig. 10. From the simulation result, we can see that when $\rho_{ZA} = 1 \times 10^{-4}$ that is smaller than 1×10^{-3} and is larger than 1×10^{-5} has the lowest steady state MSD.

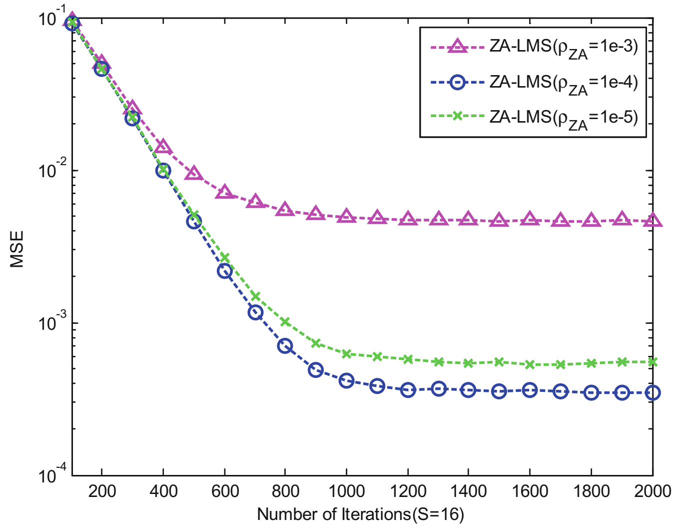

Fig. 10. ZA-LMS simulation with different ρ_{ZA} values.

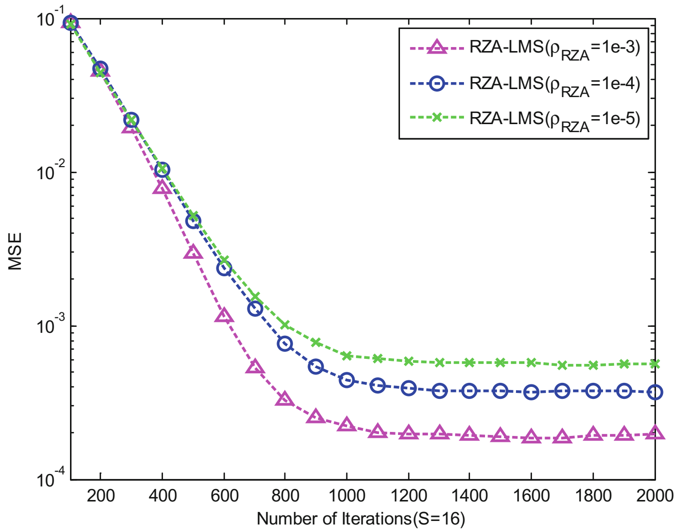

Fig. 11. RZA-LMS simulation with different ρ_{RZA} values.

The regularization parameters of another four sparse algorithm are similar as ZA-LMS. We simulate another four sparse algorithms with different regular parameters and fixed sparsity. All the simulations based on the same channel model with the length of the channel is N = 128, the sparse level of the channel is S = 16, the signal to noise ratio of the channel is SNR = 20 dB. Simulation results are obtained by taking the average of the network mean square divation (MSD) over 2000 independent Monte Carlo runs. The simulation results are shown in Figs. 11, 12, 13 and 14. The simulation results demonstrate the theoretical analysis that the smaller or larger values of ρ will degrade the performance of the algorithm.

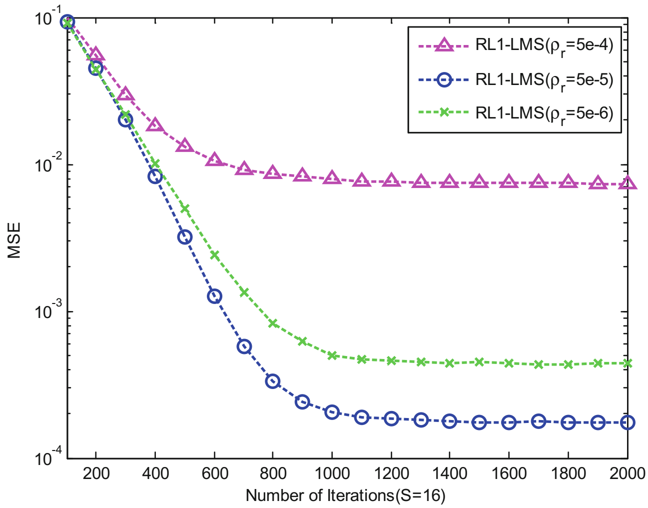

Fig. 12. RL1-LMS simulation with different $\boldsymbol{\rho_r}$ values.

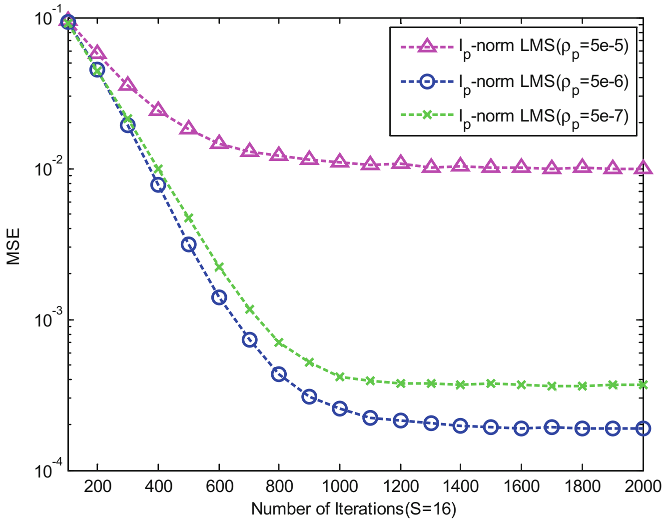

Fig. 13. ℓ_p-norm LMS simulation with different $\boldsymbol{\rho_p}$ values.

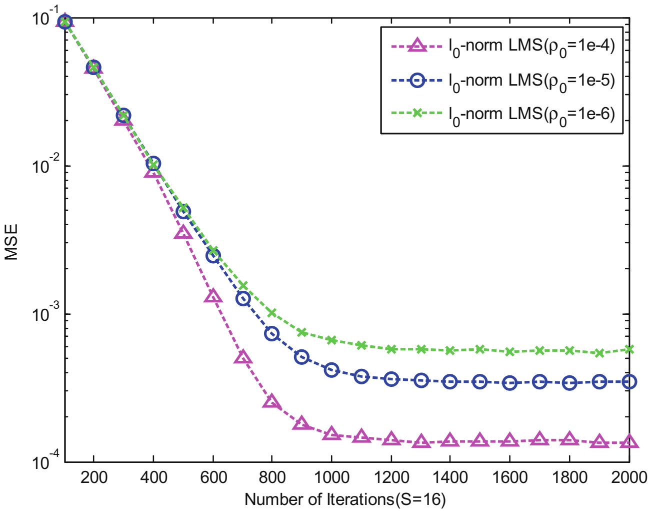

Fig. 14. ℓ_0-norm LMS simulation with different $\boldsymbol{\rho_0}$ values.

When the regularization parameter ρ_{ZA} is certain, namely, the sparse constraint is decided, the estimated solutions of the channel model with different number of zero coefficients will have various accuracy. The more the number of zero weights are, the more effectively the ZA-LMS algorithm attracts the tap-weights to zero and the vice versa. Accordingly, the performance of ZA-LMS algorithm increases with the strong sparse degree of the channel. Another four sparse algorithms have the same property. To see the influence of the spare level of channel model intuitively we simulate every algorithm with fixed regularization parameter ρ (ρ_{ZA}, ρ_{RZA}, ρ_r, ρ_p, ρ_0) and various channel model with different sparse structure. The simulation results are shown in Figs. 15, 16, 17, 18 and 19. The parameter values of the simulations are as shown in

Table 2. Parameter values of numeric simulation in Figs. 15, 16, 17, 18 and 19

Parameters	Values
Input signal	Pseudo-random binary sequences
The length of channel	N = 128
Number of dominant taps of channel vector	$S \in \{16, 32, 64\}$
Distribution of nonzero coefficients	Random Gaussian CN(0,1)
Signal to noise ratio for channel model	SNR = 20 dB
Gaussian noise distribution	Gaussian $CN\left(0, \delta_n^2\right)$
Step size	$\mu = 0.005$
Parameters of algorithms	ρ_{ZA} = 1e−4, ρ_{RZA} = 1e−3, E_{RZA} = 25, ρ_r = 5e−5, p = 0.5, ρ_p = 5e−6, E_p = 0.05, β_0 = 10, ρ_0 = 1e−4

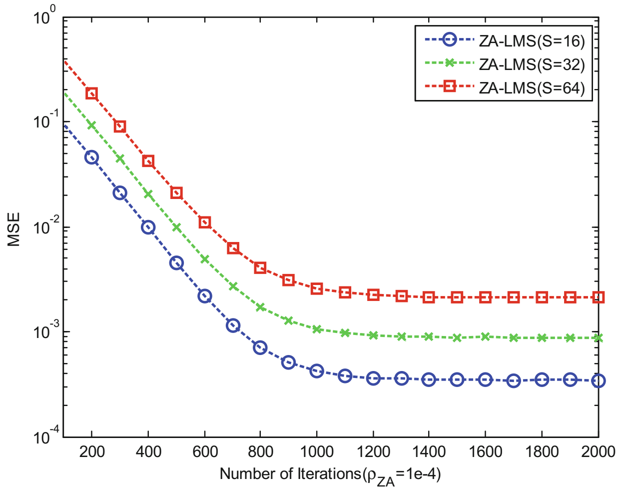

Fig. 15. Simulation of ZA-LMS algorithm with different S.

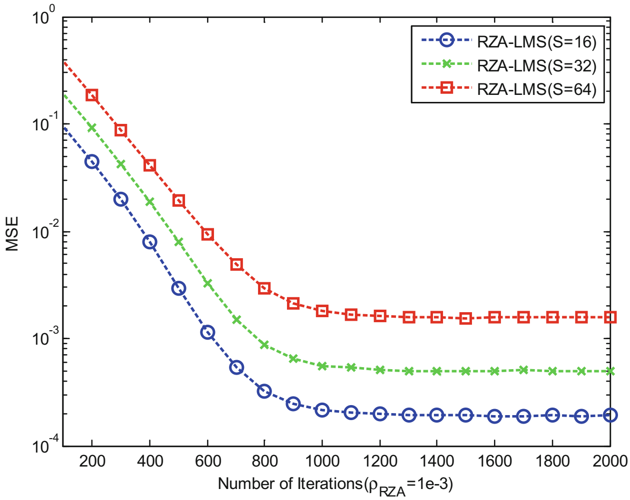

Fig. 16. Simulation of RZA-LMS algorithm with different S.

Table 2. From the simulation results we can get that the more sparse of the channel model structure, the better performance of the sparse algorithms will obtain. The conclusion is easy for us to understand because that the more numbers of the zero taps of the channel, there will be smaller bias between the estimated weights and their actual values.

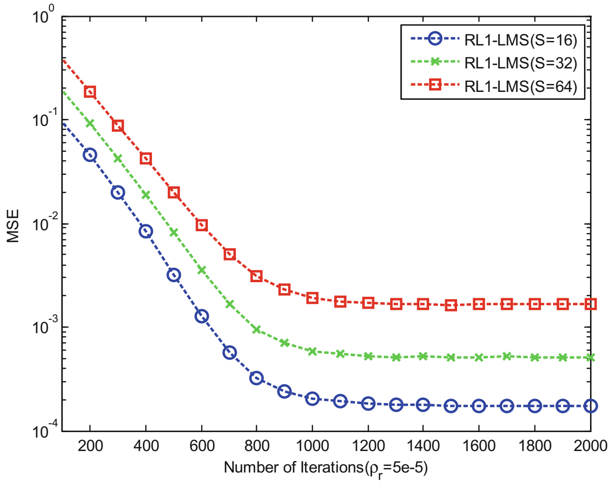

Fig. 17. Simulation of RL1-LMS algorithm with different S.

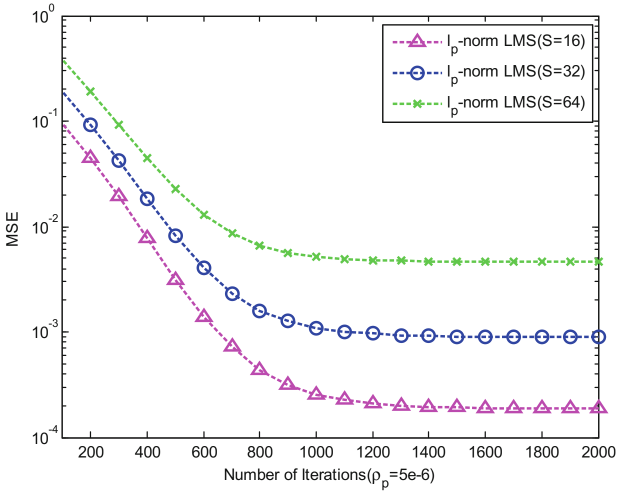

Fig. 18. Simulation of ℓ_p-norm LMS with different S.

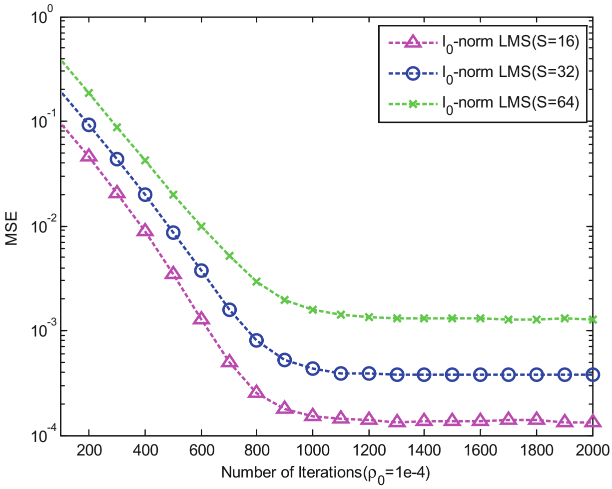

Fig. 19. Simulation of ℓ_0-norm LMS with different S.

5 Conclusions

We analyze and compare the sparse LMS algorithms performance from various aspects in this paper. The simulation results have demonstrated that the sparse LMS algorithm outperformance the traditional LMS algorithm. We know that there are many factors influence the performance of the algorithm, such as the cost function of the algorithm, the sparse constraint of the algorithm, the structure of the system model etc. To achieve the better estimation of the system, we should set proper parameter values. The more information of the system model we know, the more accurate algorithm we can design. The next work will explore more information of the system model and design more effective algorithms.

Acknowledgement. National Natural Science Foundation of China Grants (No. 61401069, No. 61701258), Jiangsu Specially Appointed Professor Grant (RK002STP16001), Innovation and Entrepreneurship of Jiangsu High-level Talent Grant (CZ0010617002), Natural Science Foundation of Jiangsu Province Grant (No. BK20170906), Natural Science Foundation of Jiangsu Higher Education Institutions Grant (No. 17KJB510044), High-Level Talent Startup Grant of Nanjing University of Posts and Telecommunications (XK0010915026) and “1311 Talent Plan” of Nanjing University of Posts and Telecommunications.

References

1. Mboup, M., Macchi, O., Bershad, N.: Steady-state superiority of LMS over LS for time-varying line enhancer in noisy environment. IEE Proc. F - Radar Sig. Process. **138**(4), 354–360 (1991)
2. Bershad, N.J., Bermudez, J.C.M., Member, S.: Stochastic analysis of the LMS algorithm for system identification with subspace inputs. IEEE Trans. Signal Process. **56**(3), 1018–1027 (2008)

3. Coleri, S., Ergen, M., Puri, A., Bahai, A.: Channel estimation techniques based on pilot arrangement in OFDM systems. IEEE Trans. Broadcast. **48**(3), 223–229 (2002)
4. Rao, H.I.K., Member, S., Farhang-boroujeny, B., Member, S.: Fast LMS/Newton algorithms for stereophonic acoustic echo cancellation. IEEE Trans. Signal Process. **57**(8), 2919–2930 (2009)
5. Taheri, O., Vorobyov, S.A.: Reweighted l_1-norm penalized LMS for sparse channel estimation and its analysis. Sign. Process. **104**, 70–79 (2014)
6. Vuokko, L., Kolmonen, V.M., Salo, J., Vainikainen, P.: Measurement of large-scale cluster power characteristics for geometric channel models. IEEE Trans. Antennas Propag. **55**(11), 3361–3365 (2007)
7. Czink, N., Yin, X., Özcelik, H., Herdin, M., Bonek, E., Fleury, B.H.: Cluster characteristics in a MIMO indoor propagation environment. IEEE Trans. Wirel. Commun. **6**(4), 1465–1474 (2007)
8. Adachi, F., Kudoh, E.: New direction of broadband wireless technology. Wirel. Commun. Mob. Comput. **7**(8), 969–983 (2007)
9. Molisch, A.F.: Ultrawideband propagation channels-theory, measurement, and modeling. IEEE Trans. Veh. Technol. **54**(5), 1528–1545 (2005)
10. Schreiber, W.F.: Advanced television systems for terrestrial broadcasting: some problems and some proposed solutions. Proc. IEEE **83**(6), 958–981 (1995)
11. Chen, Y., Gu, Y., Hero, A.O.: Sparse LMS for system identification. In: IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009, pp. 3125–3128 (2009)
12. Taheri, O., Vorobyov, S.A.: Sparse channel estimation with l_p-norm and reweighted l_1-norm penalized least mean squares. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011, pp. 2864–2867 (2011)
13. Gu, Y., Jin, J., Mei, S.: l_0-norm constraint LMS algorithm for sparse system identification. IEEE Sign. Process. Lett. **16**(9), 774–777 (2009)
14. Li, Y., Wang, Y., Jiang, T.: Low complexity norm-adaption least mean square/fourth algorithm and its applications for sparse channel estimation. In: 2016 IEEE Wireless Communications and Networking Conference Doha, Qatar, 3–6 April 2016, pp. 1–6 (2016). Electronic ISSN 1558-2612
15. Gui, G., Peng, W., Adachi, F.: Improved adaptive sparse channel estimation based on the least mean square algorithm. In: IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, 7–10 April, pp. 3105–3109 (2013)
16. Shi, K., Shi, P.: Convergence analysis of sparse LMS algorithms with l_1-norm penalty based on white input signal. Sign. Proces. **90**(12), 3289–3293 (2010)