Performance Analysis of Sparsity-Penalized LMS Algorithms in Channel Estimation

Jie Yang(✉), Hao Huang, Jie Wang, Sheng Hong, Zijian Hua, Jian Zhang, and Guan Gui(✉)

College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
{jyang, guiguan}@njupt.edu.cn

Abstract. Least mean squares (LMS) algorithm was considered as one of the effective methods in adaptive system identifications. Different from many unknown systems, LMS algorithm cannot exploit any structure characteristics. In case of sparse channels, sparse LMS algorithms are proposed to exploit channel sparsity and thus these methods can achieve better estimation performance than standard one, under the assumption of Gaussian noise environment. Specifically, several sparse constraint functions, ℓ_1-norm, reweighted ℓ_1-norm and ℓ_P-norm, are developed to take advantage of channel sparsity. By using different sparse functions, these proposed methods are termed as zero-attracting LMS (ZA-LMS), reweighted ZA-LMS (RZA-LMS), reweighted ℓ_1-norm LMS (RL1-LMS) and ℓ_p-norm LMS (LP-LMS). Our simulation results confirm the priority of the new algorithm and show that the proposed sparse algorithms are superior to the standard LMS in number scenarios.

Keywords: Gradient descent · Least mean squares · Sparse constraint Adaptive channel estimation · Compressive sensing

1 Introduction

Second-order statistical errors square based on the least mean square (LMS) algorithm has been considered one of the effective adaptive filtering methods in many applications such as channel estimation and system identification [1, 2], which is a kind of stochastic gradient algorithm. Superior to some other parameter estimation methods, e.g., recursive least squares (RLS) [3] algorithm, the LMS algorithm has the advantage that mass stochastic knowledge of the channel and the input data sequence are not required. Due to its simplicity and easy implementations, the LMS algorithm has been widely applied in signal processing and communications including system detection [4] and channel estimation [5] and so on, without considering any information about the special characteristics of the channel being estimated itself. However, due to the potential sparsity in channels [6–10], some great efforts have been made to develop such LMS algorithms that can employ the potential sparsity and achieve better parameter estimation. The method based on the idea is to add a penalty term to the cost function to perform sparse solution [11, 12]. In a typical fading communication system, the selection of the channel estimation algorithms involves the statistical information

G. Sun and S. Liu (Eds.): ADHIP 2017, LNICST 219, pp. 407–416, 2018.
https://doi.org/10.1007/978-3-319-73317-3_47

with respect to channels, the expected performance of the used algorithm and its convergence speed.

This paper is organized as follows. First we introduce the communication system model and corresponding linear adaptive algorithms. According to the given model, a standard LMS algorithm and the modifications of the LMS algorithm are provided. Particularly, the sparse channel estimation problem is considered and the sparse CIR is estimated. At last, we confirmed the effectiveness of our study.

2 System Model and Algorithms

Figure 1 shows the system model of a typical communication system in this paper. Assume that the channel vector $\boldsymbol{h} = [h_1, h_2, \ldots, h_N]^{\mathrm{T}}$ where N is the length of the CIR and $(\cdot)^{\mathrm{T}}$ denotes the transposition. $\boldsymbol{h}_k = \left[h_{1,k}, h_{2,k}, \ldots, h_{N,k}\right]^T$ denotes the estimate of the vector $\boldsymbol{h}$ at the time step k. $\boldsymbol{x}_k = [x_k, x_{k-1}, \ldots, x_{k-N+1}]^{\mathrm{T}}$ is the input data vector of the system, n_k is the additive noise at the receiver end, d_k is the actual response, $e_k = d_k - \boldsymbol{h}_k^T \boldsymbol{x}_k$ is the error signal, y_k is the system output and $\hat{y}_k$ denotes its estimate.

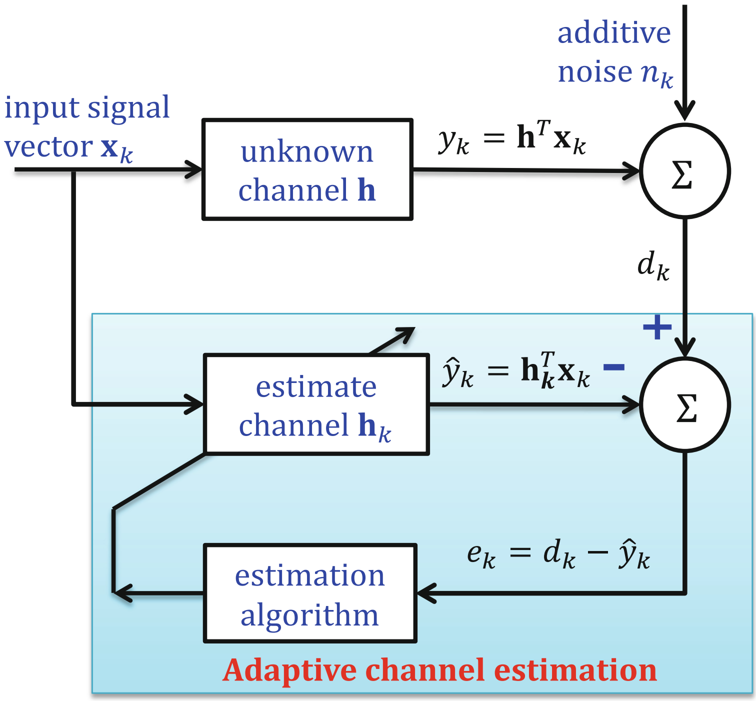


Fig. 1. Block diagram of the communication system.

2.1 Standard LMS Algorithm

Let $L_k = (1/2)e_k^2$ denotes the cost function of the standard LMS algorithm. By minimizing the cost function using the gradient descent method, the parameters of the unknown system can be identified iteratively. Therefore, the iterative equation can be given as

$$\boldsymbol{h}_{k+1} = \boldsymbol{h}_k - \mu \frac{\partial L_k}{\partial \boldsymbol{h}_k} = \boldsymbol{h}_k + \mu e_k \boldsymbol{x}_k \tag{1}$$

Here, μ is the step size which is among 0 and $\lambda_{\max}^{-1}$, where $\lambda_{\max}$ is the maximum eigenvalue of the covariance matrix of $\boldsymbol{x}_k$ (i.e., $\mathbf{R} = \mathrm{E}[\boldsymbol{x}_k \boldsymbol{x}_k^T]$), which ensures that the standard LMS algorithm converges to the optimum point.

2.2 ZA-LMS Algorithm

If most of the coefficients in the vector $\boldsymbol{h}$ are zeros or insignificant values, then the CIR is called sparse channel. In this case, the l_1-norm of $\boldsymbol{h}_k$ can be used to penalize the non-sparse solutions. Add it to the standard LMS cost function and then we can get the new cost function $L_k^{ZA} = (1/2)e_k^2 + \gamma_{ZA}\|\boldsymbol{h}_k\|_{l_1}$, where $\|.\|_{l_1}$ denotes the l_1-norm of a vector and γ_{ZA} is a corresponding weight for the penalty term. It's remarkable that the new cost function is convex, so that the gradient descent method can be guaranteed to be convergent under some conditions. The corresponding algorithm is called the zero attracting LMS (ZA-LMS) and its iterative formula is

$$\boldsymbol{h}_{k+1} = \boldsymbol{h}_k + \mu e_k \boldsymbol{x}_k - \rho_{ZA} sgn(\boldsymbol{h}_k) \tag{2}$$

where $\rho_{ZA} = \mu\gamma_{ZA}$ and signum function sgn(·) is denoted as 0 for x = 0, 1 for x > 0, and −1 for x < 0 and $sgn(\boldsymbol{h}_k)$ is the sparse penalty strength of the ZA-LMS.

2.3 RZA-LMS Algorithm

To take more advantage of the sparsity of the channel, we can use the l_0-norm to penalize the non-sparse solutions. However, since l_0-norm penalty has very high computation complexity, a approximate penalty is introduced. And then the cost function becomes

$$L_k^{RZA} = \left(\frac{1}{2}\right)e_k^2 + \gamma_{RZA} \sum\nolimits_{i=1}^{N} \log(1 + \frac{\mathbf{h}_{ki}}{\in'_{RZA}}) \tag{3}$$

where $\mathbf{h}_{ki}$ is the i-th entry of the channel weights $\boldsymbol{h}_k$. γ_{RZA} and $\in'_{RZA}$ are some positive numbers. Since the logarithmic constraint in (3) that resembles the l_0-norm penalty can describe the sparse channel more accurate, it is expected that the corresponding algorithm which is defined as the reweighted ZA-LMS (RZA-LMS) will gain a more accurate estimation than the ZA-LMS. The iterative formula of the corresponding algorithm is

$$\boldsymbol{h}_{k+1} = \boldsymbol{h}_k + \mu e_k \boldsymbol{x}_k - \rho_{RZA} \frac{sgn(\boldsymbol{h}_k)}{1 + \in_{RZA} |\boldsymbol{h}_k|} \tag{4}$$

where $\rho_{RZA} = \mu \gamma_{RZA} \in_{RZA}$, $\in_{RZA} = 1/\in'_{RZA}$, absolute value $|\cdot|$, and $\frac{sgn(\boldsymbol{h}_k)}{1+\in_{RZA}|\boldsymbol{h}_k|}$ is the sparse penalty strength of the RZA-LMS.

2.4 LP- LMS Algorithm

In order to further obtain sparse information, p-norm (where p is among 0 and 1) spare function is adopted in LMS-type channel estimation. We called it as for LP-LMS algorithm. The new function is more close to the l_0-norm and as the value of p becomes smaller, it resembles the l_0-norm more. Thus, the cost function of LP-LMS algorithm is given as

$$L_k^{l_p} = \left(\frac{1}{2}\right) e_k^2 + \gamma_p ||\boldsymbol{h}_k||_{l_p} \tag{5}$$

where $||.||_{l_p}$ denotes the l_p-norm of the vector and γ_p denotes the corresponding weight term. It is notice that the cost function (5) is nonconvex and the analysis of the global convergence and consistency of the corresponding algorithm is problematic. However, as it will be seen in the next section, the method based on (5) shows better performance than the RZA-LMS which faces the same problems. Using gradient descent, the update equation based on (5) can be derived as

$$\boldsymbol{h}_{k+1} = \boldsymbol{h}_k + \mu e_k \boldsymbol{x}_k - \rho_p \frac{\left(||\boldsymbol{h}_k||_p\right)^{1-p} sgn(\boldsymbol{h}_k)}{\in_p + |\boldsymbol{h}_k|^{1-p}} \tag{6}$$

where $\rho_p = \mu \gamma_p$, $\in_p$ is some number near to zero and $\frac{\left(||\boldsymbol{h}_k||_p\right)^{1-p} sgn(\boldsymbol{h}_k)}{\in_p + |\boldsymbol{h}_k|^{1-p}}$ is the sparse penalty strength of the l_p-norm penalized LMS.

2.5 RL1-LMS Algorithm

One of alternative way to exploit channel sparsity by using RL1 penalty in accordance with mean square error term. This method considers a penalty term proportional to the reweighted l_1-norm of the coefficient vector. Compared to the standard l_1-norm minimization, this method can get better channel estimation performance. The mentioned cost function above can be written as

$$L_k^{rl1} = \left(\frac{1}{2}\right) e_k^2 + \gamma_r ||\boldsymbol{s}_k \boldsymbol{h}_k||_{l_1} \tag{7}$$

where γ_r is a tradeoff parameter and RL1 row vector $\boldsymbol{s}_k$ are given as

$$[\boldsymbol{s}_k]_i = \frac{1}{\in_r + \left|[\boldsymbol{h}_{k-1}]_i\right|}, i = 1, \ldots, N \tag{8}$$

with small positive parameter $\in_r$. Hence, the RL1-LMS algorithm is derived as

$$\boldsymbol{h}_{k+1} = \boldsymbol{h}_k + \mu e_k \boldsymbol{x}_k - \rho_r \frac{sgn(\boldsymbol{h}_k)}{\in_r + |\boldsymbol{h}_{k-1}|} \tag{9}$$

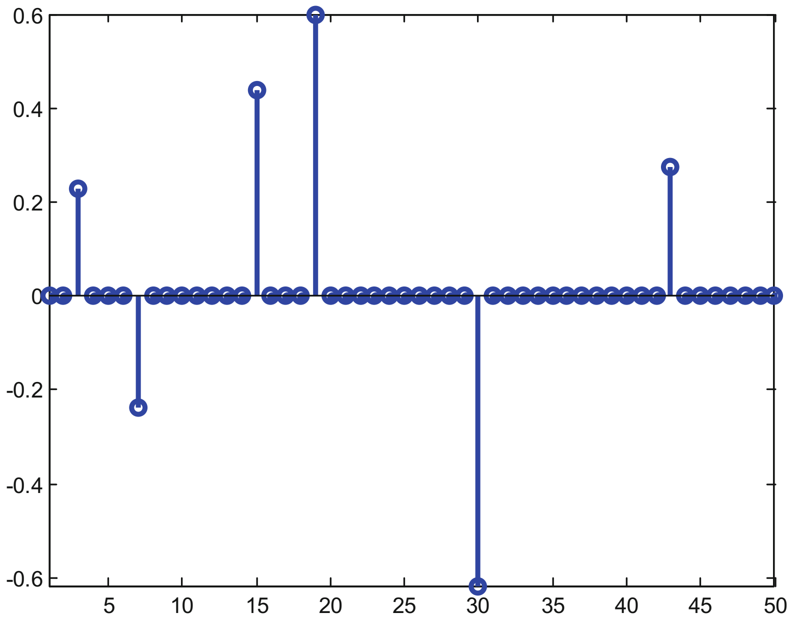
where $\rho_r = \mu\gamma_r$ and $\frac{sgn(\boldsymbol{h}_k)}{\in_r + |\boldsymbol{h}_{k-1}|}$ is the sparse penalty strength of the reweighted l_1-norm penalized LMS.

3 Simulation Results

Compared with the standard LMS algorithm, other modified LMS algorithms take the sparsity of the CIR into account. Figure 2(a) is a sparse vector (the number of non-zero values is much smaller than the total length of the vector) diagram. Figure 2(b) shows the sparse penalty strengths for the algorithms tested versus the coefficient component of the estimate $\boldsymbol{h}_k$ of the vector $\boldsymbol{h}$ at the time step k where the CIR is assumed to $\boldsymbol{h_k} = [-1 : 0.001 : 1]$ for all algorithms and p is set to 0.5 in the ℓ_p-norm penalized method. For the ZA-LMS, the sparse penalty strength is zero at the position of zero and is the value of 1 at the non-zero position. Therefore, when the sparse channel vector is disturbed by noise, the value of the sparse position may fluctuate near the value of 0 and the ZA-LMS algorithm can result in obvious errors. However, for the RZA-LMS, the ℓ_p-norm penalized LMS and reweighted ℓ_1-norm penalized LMS, the closer to the value of zero the value of the sparse channel vector coefficient is, the greater the sparse penalty strength is and the higher the probability of taking zero is; the farther away from the value of zero the value of the sparse channel vector coefficient is, the smaller the sparse penalty strength is and the lower the probability of taking zero is. Overall, the sparse penalty strength of the ℓ_p-norm penalized LMS is greater than that of the reweighted ℓ_1-norm penalized LMS and the sparse penalty strength of the reweighted ℓ_1-norm penalized LMS is greater than that of the RZA-LMS.

As is shown in Fig. 3, ZA-LMS, RZA-LMS, the ℓ_p-norm penalized LMS and the reweighted ℓ_1-norm penalized LMS take different regularization parameters to obtain simulation results of MSEs in contrast to the number of iterations respectively in the other same conditions. The step size is set to $\mu = 0.05$ and the signal-to-noise ratio (SNR) is set to 10 dB, which implies that the MSEs are averaged at 2000 simulations. The length of the CIR is 16 and the sparsity level is set to 1, which means that there is only one nonzero tap in the CIR, but the nonzero position is allocated randomly. The other parameters are set to $\in_{\mathrm{RZA}} = 10$, $\in_{l\mathrm{p}} = \in_{rl1} = 0.05$, $p = 0.5$. Figure 3 shows the convergence speed and the steady state MSE is related to ρ and ρ is larger, the convergence speed is faster but the MSE is also larger at steady state. With ρ decreasing, the MSE of the steady state decreases first, then increases. The minimum steady state MSEs of ZA-LMS, RZA-LMS, the ℓ_p-norm penalized LMS and the reweighted ℓ_1-norm penalized LMS appear in $\rho = 10^{-3}, \rho = 10^{-2}, \rho = 10^{-1}, \rho = 10^{0}$, respectively.

(a)



(b)

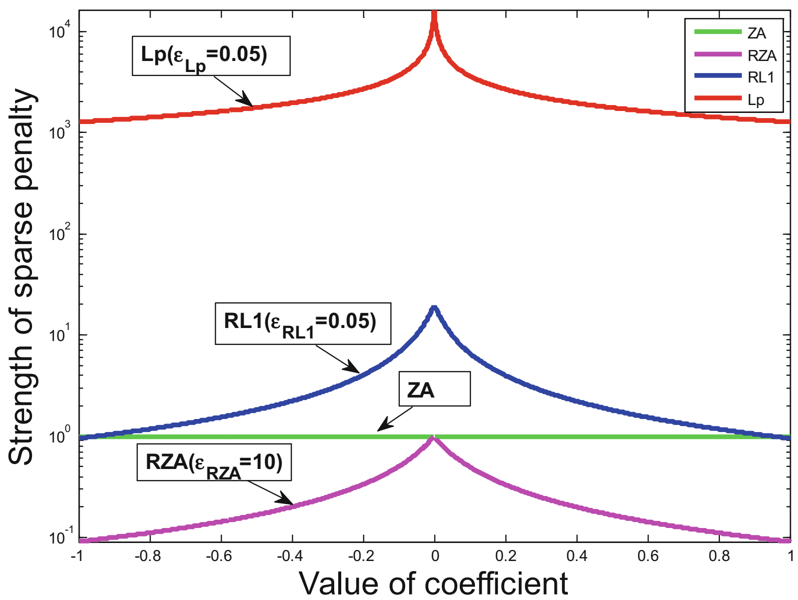


Fig. 2. Strengths of sparse penalty of different estimation algorithms vs number of iterations.

(a)

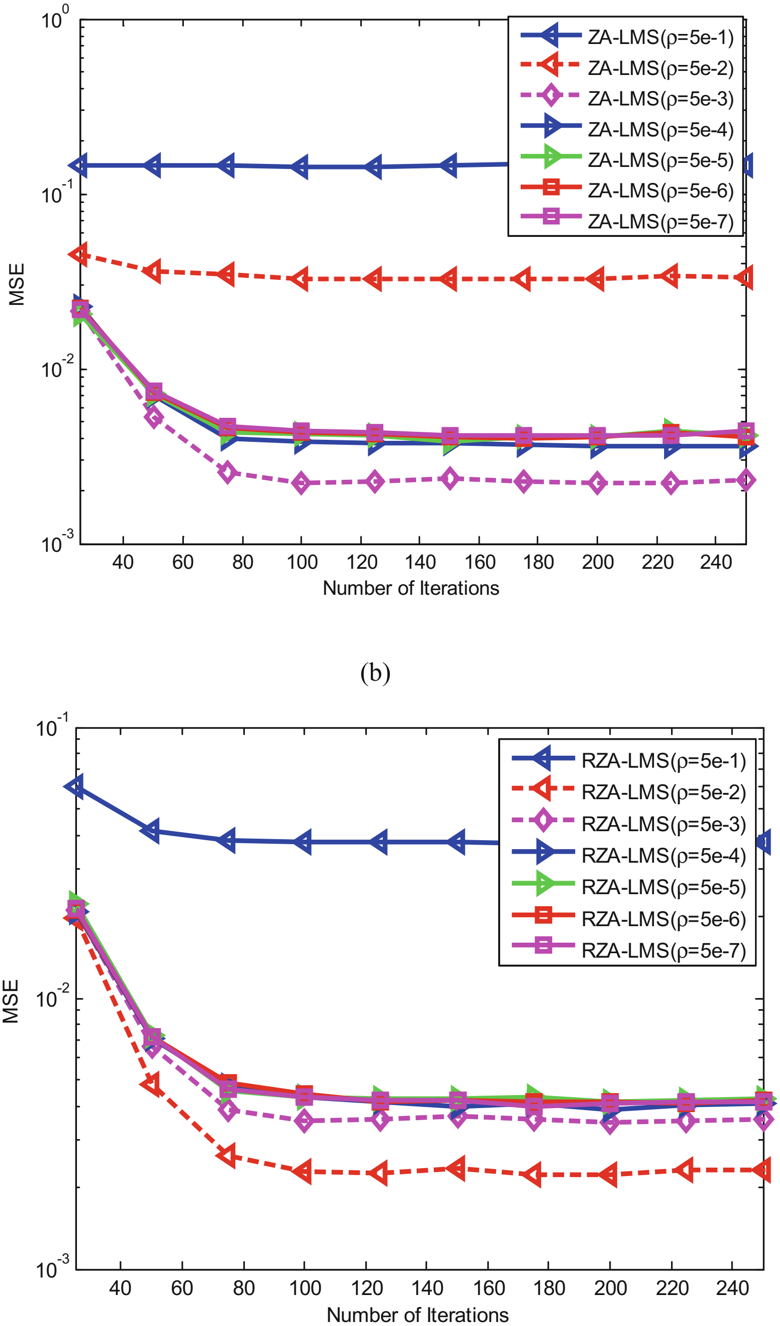
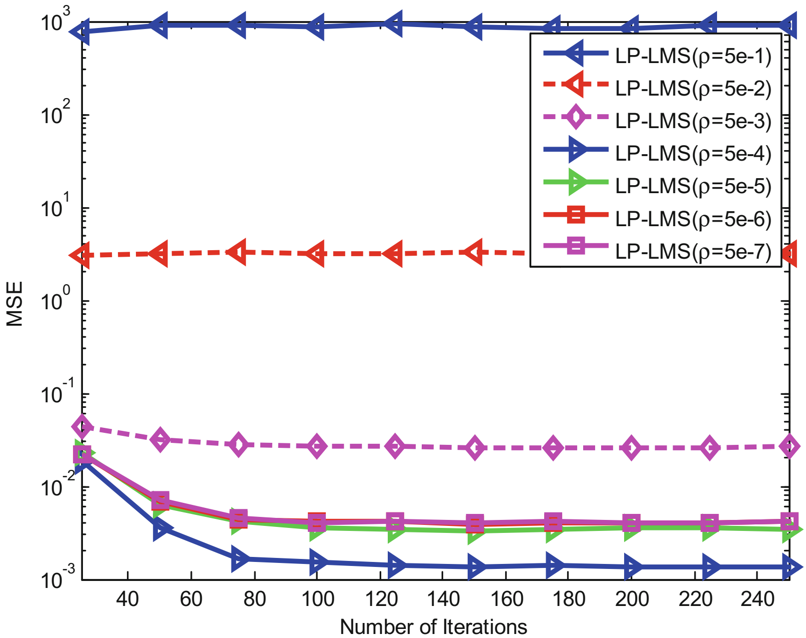


Fig. 3. MSE comparisons with respect to iterations in SNR = 10 dB.

(c)



(d)

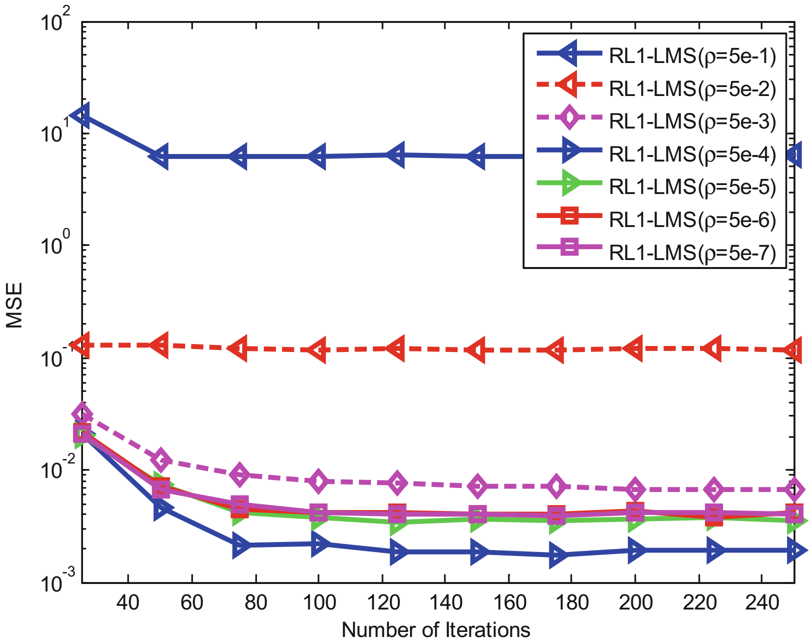


Fig. 3. (*continued*)

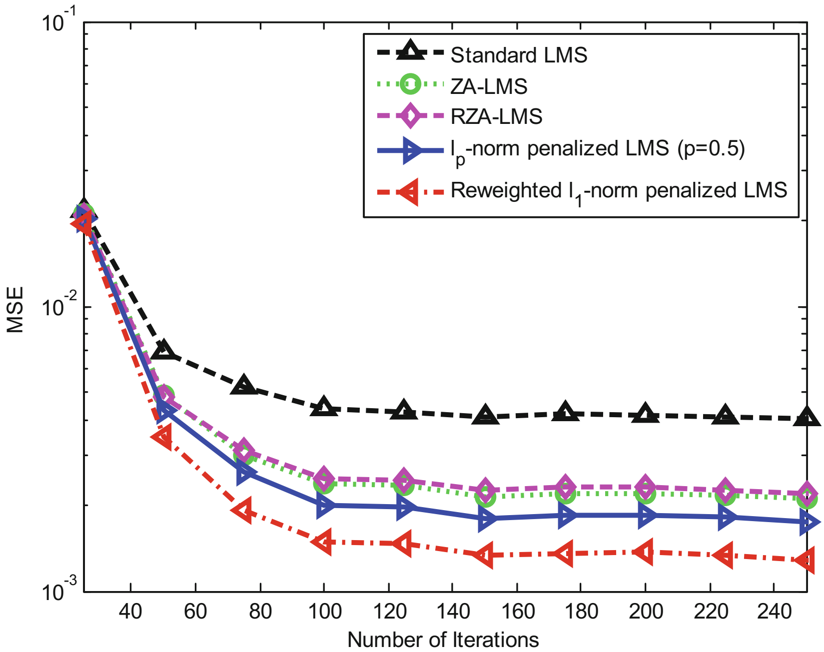


Fig. 4. MSE comparisons with respect to iterations in SNR = 20 dB.

Figure 4 indicates the MSEs of different estimation algorithms with respect to iterations in SNR = 10 dB. The performance of the improved sparse LMS algorithms is compared to that of the standard LMS. The step size is set to $\mu = 0.05$, the signal-to-noise ratio (SNR) is set to 20 dB, the channel length of the CIR is 16, the sparsity level is set to 1 and number of iterations is 2000 times for all LMS algorithms. The other parameters are set to $\in_{\mathrm{RZA}} = 10, \in_{lp} = \in_{rl1} = 0.05, p = 0.5$. This can be observed by observing Fig. 4. It is worth noting that both ZA-LMS and RZA-LMS algorithms demonstrate very close performance, while they are much better than LMS. In case of the MSE curves shown in the Fig. 4, it can also be concluded that the reweighted ℓ_1-norm penalized LMS has better performance than the ℓ_p-norm penalized LMS. By examining Fig. 4, it can also be seen that both the reweighted l1-norm and the ℓ_p-norm penalized LMS algorithms have better performance than ZA-LMS and RZA-LMS algorithms.

4 Conclusions

This paper considers the sparsity of the communication system and applies the sparisty to channel estimation with LMS algorithms. Quantitative simulations and analysis indicates that the improved LMS algorithms outperform the standard LMS algorithm with regard to sparse CIR. In addition, for the RZA-LMS, the ℓ_p-norm penalized LMS and reweighted ℓ_1-norm penalized LMS, the closer to the value of zero the value of the sparse channel vector coefficient is, the greater the sparse penalty strength is and the higher the probability of taking zero is and vice versa, which can refrain from the sparse channel vector disturbed by noise that can make the value of the sparse position fluctuate near the value of zero and cause great errors. Compared to the ZA-LMS and RZA-LMS, the ℓ_p-norm penalized LMS and reweighted ℓ_1-norm penalized LMS have better performance in simulation results.

References

1. Chen, S., Kao, Y., Tsai, K.: A new efficient LMS adaptive filtering algorithm. IEEE Trans. Circ. Syst. Analog Digit. Sig. Process. **43**(5), 372–378 (1996)
2. Godavarti, M., Hero, A.O.: Partial update LMS algorithms. IEEE Trans. Sig. Process. **53**(7), 2382–2399 (2005)
3. Stanciu, C., Udrea, M., Anghel, C.: Improved regularization for a low-complexity RLS algorithm. In: IEEE TELFOR, vol. 2, no. 5, pp. 2–5 (2016)
4. Bershad, N.J., Bermudez, J.C.M., Member, S.: Stochastic analysis of the LMS algorithm for system identification with subspace inputs. IEEE Trans. Sig. Process. **56**(3), 1018–1027 (2008)
5. Rana, M.: Performance comparison of LMS and RLS channel estimation algorithms for 4G MIMO OFDM systems. In: IEEE Proceedings of 14th International Conference on Computer and Information Technology, ICCIT, pp. 22–24 (2011)
6. Truhachev, D., Schlegel, C., Yin, D.: The impact of filtering on the resolution of sparsity in channel estimation for OFDM systems. IEEE Ultra Maritime Digital Communications Centre, no. 1 (2016)
7. Zhang, J., Zhang, Y., Yawei, Y., Ruijie, X., Zheng, Q., Zhang, P.: 3D MIMO: how much does it meet our expectation observed from channel measurements? IEEE J. Sel. Areas Commun. **35**(8), 1887–1903 (2017)
8. Zhang, J., Tang, P., Tian, L., Hu, Z., Wang, T., Haiming, W.: 6–100 GHz research progress and challenges for fifth generation (5G) and future wireless communication from channel perspective. SCIENCE CHINA Inf. Sci. **60**(8), 1–16 (2017)
9. Rappaport, T.S., Xing, Y., MacCartney Jr., G.R., Molisch, A., Mellios, E., Zhang, J.: Overview of millimeter wave channel models for fifth-generation wireless networks. IEEE Trans. Antennas Propag. **PP**(99), 1 (2017)
10. Zhang, J., Pan, C., Pei, F., Liu, G., Chen, X.: Three-dimensional fading channel models: a survey of elevation angle research. IEEE Commun. Mag. **52**(6), 218–226 (2014)
11. Taheri, O., Vorobyov, S.A.: Reweighted L1-norm penalized LMS for sparse channel estimation and its analysis. Sig. Process. **104**, 70–79 (2014)
12. Taheri, O., Vorobyov, S.A.: Sparse channel estimation with Lp-norm and reweighted L1-norm penalized least mean squares. In: IEEE ICASSP, 22–27 May 2011, pp. 2864–2867 (2011)