Regularization Selection Method for LMS-Type Sparse Multipath Channel Estimation

Zhengxing Huang⋆, Guan Gui †, Anmin Huang ‡, Dong Xiang⋆, and Fumiyki Adachi †
⋆ Department of Software Engineering, Tsinghua University, Beijing, China
† Department of Communication Engineering, Tohoku University, Sendai, Japan
‡ Department of Electronics and Information Engineering, Jinggangshan University, Jian, China

***Abstract*—Least mean square (LMS)-type adaptive sparse algorithms have been attracting much attention on sparse multipath channel estimation (SMPC) due to their two advantages: low computational complexity and reliability. By introducing ℓ_1-norm sparse constraint function into LMS algorithm, both *zero-attracting least mean square* (ZA-LMS) and *reweighted zero-attracting least mean square* (RZA-LMS) have been proposed for SMPC. It is well known that the performance of the SMPC is decided by *regularization parameter* which balances channel estimation error and sparse penalty strength. However, optimal regularization parameter selection has not yet considered in the two proposed algorithms. Based on the compressive sensing theory, in this paper, we explain the mathematical relationship between Lasso and LMS-type adaptive sparse algorithms. Later, an approximate optimal regulation parameter selection method is proposed for ZA-LMS and RZA-LMS, respectively. Monte Carlo based computer simulations are presented to show the effectiveness of our propose method.**

Keywords—regularization parameter selection, least mean square (LMS); adaptive sparse channel estimation; zero-attracting least mean square (ZA-LMS); reweighted zero-attracting least mean square (RZA-LMS).

I. INTRODUCTION

The demand for high-speed data services is getting more insatiable due to the number of wireless subscribers roaring increase in the next generation wireless communication systems. Various portable wireless devices, e.g., smart phones and laptops, have generated rising massive data traffic [1]. It is well known that the broadband transmission is an indispensable technique for realizing Gigabit wireless communication [2][3]. However, the broadband signal is susceptible to interference by frequency-selective channel fading. In the sequel, the broadband channel is described by a sparse channel model in which multipath taps are widely separated in time, thereby create a large delay spread [4]. In other words, unknown channel impulse response (CIR) in broadband wireless communication system is often described by sparse channel model, supporting by a few large coefficients. In other words, most of channel coefficients are zero or close to zero while only a few channel coefficients are dominant (large value) to support the channel. A typical example of sparse channel is shown in Fig. 1, where the number of dominant channel taps is 4 while the length of channel is 16.

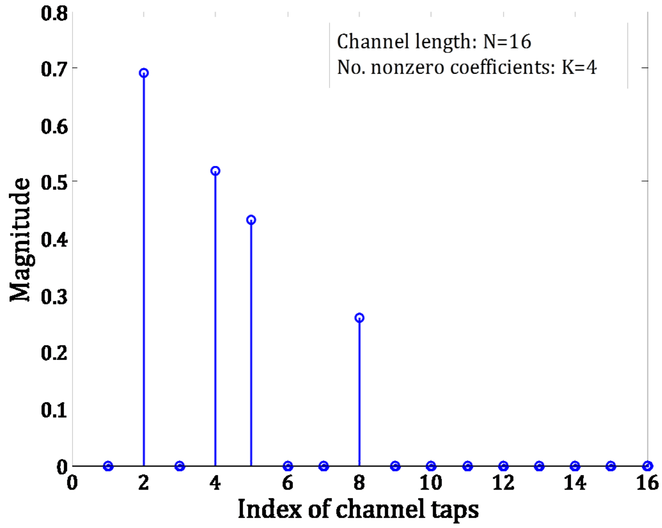


Fig. 1. A typical example of sparse multipath channel.

Traditional least mean square (LMS) algorithm is one of the most popular methods for adaptive system identification [5], e.g. channel estimation. Indeed, LMS-based adaptive channel estimation can be easily implemented by LMS-based filter due to its low computational complexity or fast convergence speed. However, the standard LMS-based method never takes advantage of channel sparse structure as prior information and then it may loss some estimation performance.

Recently, many algorithms have been proposed to take advantage of sparse structure of the channel. For example, based on the theory of compressive sensing (CS) [6], [7], various sparse channel estimation methods have been proposed in [8–13]. For one thing, these CS-based sparse channel estimation methods require that the training signal matrices satisfy the restricted isometry property (RIP) [14]. However, design these kinds of training matrices is non-deterministic polynomial-time (NP) hard problem [21]. For another thing, some of these methods achieve robust estimation at the cost of high computational complexity, e.g., sparse channel estimation using least-absolute shrinkage and selection operator (LASSO) [15]. To avoid the high computational complexity on sparse channel estimation, a variation of the LMS algorithm with ℓ_1-norm penalty term in the LMS cost function has also been developed in [16], [17]. The ℓ_1-norm penalty was incorporated into the cost function of conventional LMS algorithm, which resulted in two sparse LMS algorithms, namely *zero-attracting least mean square* (ZA-LMS) and *reweighted zero-attracting least mean square* (RZA-LMS) [16]. Moreover, improved adaptive sparse channel estimators were proposed in [17–19]. It was well known that adaptive sparse channel estimation methods depend on regularization parameter which controls estimation error and channel sparsity. As the authors best understanding, however, there is no paper reported that regularization parameter selection method for ZA-LMS and RZA-LMS.

In this paper, we propose a regularization parameter selection method for achieving optimal sparse LMS channel estimation in different signal-to-noise ratio (SNR) regimes.

The remainder of this paper is organized as follows. Section II introduces sparse system model. Section III reviews LMS-type adaptive sparse channel estimation methods and presents problem formulation. In section V, we propose Monte Carlo-based regularization selection method using different simulation results. Concluding remarks are presented in Section V.

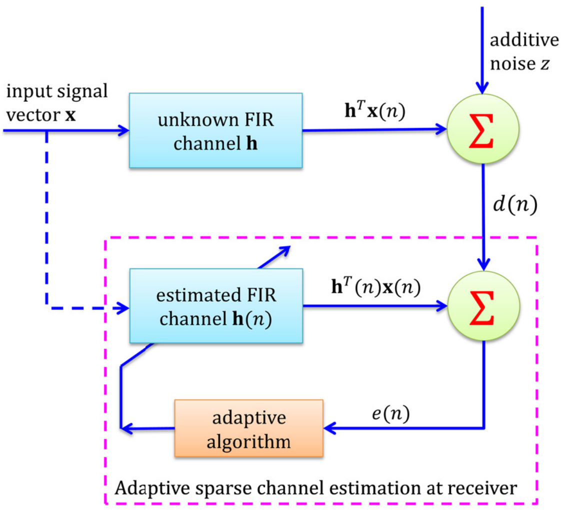


Fig.2. An adaptive sparse channel estimation based sparse multipath communication system.

II. System Model

Consider a sparse multipath adaptive communication system, as shown in Fig. 2. The input signal $\mathbf{x}(t)$ and ideal output signal $d(t)$ are related by

$$d(t) = \mathbf{h}^T\mathbf{x}(t) + z(t), \tag{1}$$

where $\mathbf{h} = [h_0, h_1, \dots, h_{N-1}]^T$ is a N-length sparse channel vector which is supported only by K dominant channel taps, $\mathbf{x}(t) = [x(t), x(t-1), \dots, x(t-N+1)]^T$ is N-length input signal vector and $z(t)$ is an additive noise variable at time t. The objective of LMS adaptive filter is to estimate the unknown sparse channel coefficients $\mathbf{h}$ using the input signal $\mathbf{x}(t)$ and ideal output signal $d(t)$. n-th adaptive estimation error is termed as $e(n)$. For a better understanding, input signal $\mathbf{x}(t)$ and output signal $d(t)$ are also revised as $\mathbf{x}(n)$ and $d(n)$, respectively, where n denotes adaptive iterative times. At the time t, please note that both $\mathbf{x}(n)$ and $d(n)$ are invariant. According to Eq. (1), channel estimation error $e(n)$ is written as

$$e(n) = d(n) - \mathbf{h}^T(n)\mathbf{x}(n), \tag{2}$$

where $\mathbf{h}(n)$ is the LMS adaptive channel estimator. Based on Eq. (2), LMS cost function can be given by

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

Hence, the update equation of LMS adaptive channel estimation is derived by

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu e(n)\mathbf{x}(n), \tag{4}$$

where $\mu \in (0, 2/\gamma_{\max})$ is a step size of gradient descend step-size and $\gamma_{\max}$ is the maximum eigenvalue of the covariance matrix of $\mathbf{x}(n)$.

III. LMS-type Adaptive Sparse Channel Estimation Methods

From the above Eq. (4), we can find that the LMS-based channel estimation method never take advantage of sparse structure in $\mathbf{h}$. The standard LMS-based channel estimation can be concluded as

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \text{adaptive update}. \tag{5}$$

Unlike the standard LMS method in (5), channel sparsity can be exploited by introducing ℓ_1-norm penalty to LMS-type cost function [16], [17]. Hence, the LMS-based adaptive sparse channel estimation can be written as

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \text{adaptive update} + \text{sparse penalty}. \tag{6}$$

From above update Eq. (6), the objective of adaptive sparse channel estimation is introducing different sparse penalties to take the advantage of sparse structure as for prior information.

A. ZA-LMS algorithm

To exploit the channel sparsity in CIR, the cost function of ZA-LMS [16] is given by

$$L_{ZA}(n) = \frac{1}{2}e^2(n) + \lambda_{ZA}\|\mathbf{h}(n)\|_1, \tag{7}$$

where $\lambda_{ZA} = \rho_{ZA}\sigma_n\sqrt{2N}/100$ is a regularization parameter which balances the adaptive estimation error and sparse penalty of $\mathbf{h}(n)$. Please note that the ρ_{ZA} is a setting parameter which controls the λ_{ZA}. The corresponding update equation of ZA-LMS was written as

$$\mathbf{h}(n+1) = \mathbf{h}(n) - \mu\frac{\partial L_{ZA}(n)}{\partial \mathbf{h}(n)}$$

$$= \mathbf{h}(n) + \mu e(n)\mathbf{x}(n) - \kappa_{ZA}\text{sgn}\{\mathbf{h}(n), \tag{8}$$

where $\kappa_{ZA} = \mu\lambda_{ZA}$ and $\text{sgn}\{\cdot\}$ is a component-wise function which is defined as

$$\text{sgn}(h) = \begin{cases} h/|h|, & \text{when } h \neq 0 \\ 0, & \text{when } h = 0 \end{cases}, \tag{9}$$

where the h is one of channel taps of $\mathbf{h}$. From the update equation in Eq. (8), the second term attracts the small filter coefficients to zero, which speed up convergence when the most of the channel coefficients **h** are zeros. Here, the sparse penalty function in Eq. (8) is defined as

$$G_{ZA}(\mathbf{h}(n)) = \mathrm{sgn}(\mathbf{h}(n)), \quad (10)$$

which is depicted as shown in Fig. 3.

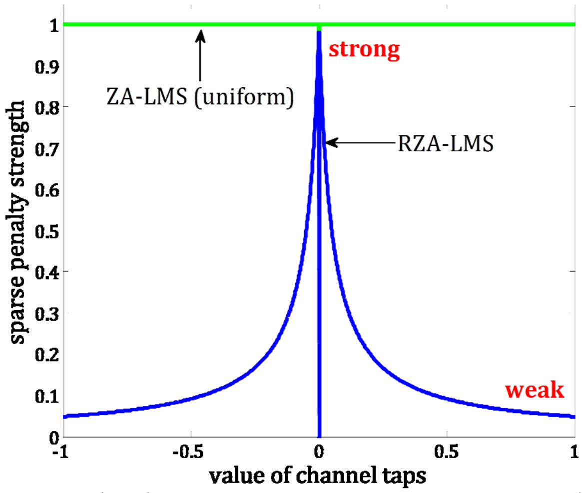


Fig. 3. $G_{ZA}(\mathbf{h}(n))$ for different channel taps is uniform while $G_{RZA}(\mathbf{h}(n))$ is strong for small channel taps and weak for big channel taps.

B. *RZA-LMS algorithm*

The ZA-LMS cannot distinguish between zero taps and non-zero taps since all the taps are forced to zero uniformly; therefore, its performance will degrade in less sparse systems. Motivated by reweighted ℓ_1-minimization sparse recovery algorithm [20], adaptive sparse channel estimation using zero-attracting least mean square (RZA-LMS) was proposed in . The cost function of RZA-LMS is given by

$$L_{RZA}(n) = \frac{1}{2}e^2(n) + \lambda_{RZA}\sum_{i=1}^{N}\log(1 + \varepsilon_{RZA}|h_i|), \quad (11)$$

where $\lambda_{RZA} = \rho_{RZA}\sigma_n\sqrt{2N}$ is a regularization parameter which trades off the estimation error and channel sparsity. It was worth note that the ρ_{RZA} is a setting parameter which controls the λ_{RZA}. According to Eq. (11), the corresponding update equation was given by

$$\begin{aligned}\mathbf{h}(n+1) &= \mathbf{h}(n) - \mu\frac{\partial L_{RZA}(n)}{\partial \mathbf{h}(n)} \\ &= \mathbf{h}(n) + \mu e(n)\mathbf{x}(n) \\ &\quad -\mu\lambda_{RZA}\varepsilon_{RZA}\sum_{i=1}^{N}\frac{\mathrm{sgn}(|h_i(n)|)}{1+\varepsilon_{RZA}|h_i(n)|} \\ &= \mathbf{h}(n) + \mu e(n)\mathbf{x}(n) - \kappa_{RZA}\frac{\mathrm{sgn}(\mathbf{h}(n))}{1+\varepsilon_{RZA}|\mathbf{h}(n)|}, \quad (12)\end{aligned}$$

where $\kappa_{RZA} = \mu\lambda_{RZA}\varepsilon_{RZA}$ is a parameter which depends on step-size μ, regularization parameter λ_{RZA} and threshold parameter ε_{RZA}, respectively. In Eq. (12), if magnitudes of $h_i(n)$, $i = 1,2,\ldots,N$ are smaller than $1/\varepsilon_{RZA}$, then these channel coefficients will be replaced by zeros in high probability. Here, the sparse penalty function in Eq. (12) is defined as

$$G_{RZA}(\mathbf{h}(n)) = \frac{\mathrm{sgn}(\mathbf{h}(n))}{1+\varepsilon_{RZA}|\mathbf{h}(n)|}. \quad (13)$$

Take $\varepsilon_{RZA} = 20$ as for an example, sparse penalty function $G_{\mathrm{RZA}}(\mathbf{h}(n))$ in Eq. (13) can be depicted as in Fig. 3.

IV. COMPUTER SIMULATIONS

In this section, we compare the performance of proposed channel estimators using 1000independent Monte-Carlo runs for averaging. The length of sparse multipath channel **h** is set as $N = 16$ and its number of dominant taps is set as $K = 2$ and 4 respectively. The values of dominant channel taps follow random Gaussian distribution and the positions of dominant taps are randomly allocated within the length of **h** which is subjected to $E\{\|\mathbf{h}\|_2^2\} = 1$. The signal-to-noise ratio (SNR) is defined as $10\log(E_0/\sigma_n^2)$, where E_0 is transmitted power. Here, we set the SNR range from 5dB to 30dB. Simulation parameters are listed in Tab. I.

TABLE I. SIMULATION PARAMETERS FOR LMS-BASED ADAPTIVE SPARSE CHANNEL ESTIMATION.

Type of parameters	Value
Step-size μ	5e-2
Channel length	$N = 16$
Number of nonzero taps	2 & 4
Channel distribution	Random Gaussian

The estimation performance is evaluated by mean square deviation (MSD) standard which is defined as

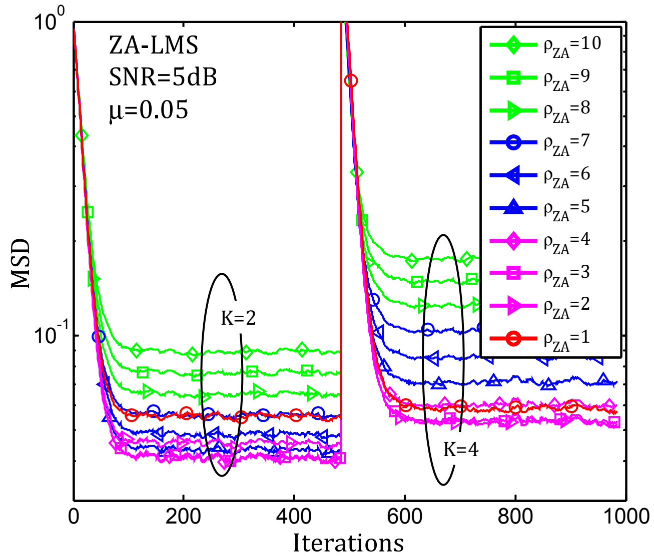
$$\mathrm{MSD}(\mathbf{h}(n)) = \mathrm{E}\{\|\mathbf{h}(n) - \mathbf{h}\|_2^2\}, \quad (14)$$

where $\mathrm{E}[\cdot]$ denotes expectation operator, **h** and $\mathbf{h}(n)$ are the actual channel vector and its estimator, respectively.

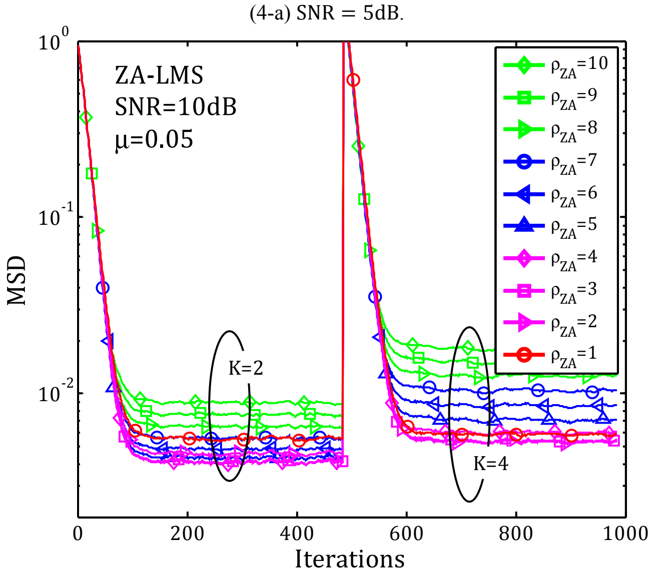
The regularization parameter of ZA-LMS is denoted by $\lambda_{ZA} = \rho_{ZA}\sigma_n\sqrt{2N}/100$. Since noise variance σ_n^2 and channel length N are given by the system, hence, λ_{ZA} depends on the parameter ρ_{ZA}, that is $\lambda_{ZA}\sim\mathcal{O}(\rho_{ZA})$. We evaluate ZA-LMS based adaptive sparse channel estimation method with different SNRs as shown in Fig. 4(a-f). Different MSD estimation performance curves are depicted as different parameters ρ_{ZA}. In Fig. 4(a), ZA-LMS can achieve approximate optimal performance using parameter $\rho_{ZA} = 3$ than previous method using other parameters at the SNR = 5dB. As the SNR increasing, ZA-LMS can also achieve the approximate optimal sparse channel estimation. According to six sub-figures in Fig. 4, $\rho_{ZA} = 3$ is chosen as approximate optimal regularization parameter for ZA-LMS.

The regularization parameter of RZA-LMS is denoted as $\lambda_{RZA} = \rho_{RZA}\sigma_n\sqrt{2N}$. Hence, λ_{RZA} depends on the parameter ρ_{RZA}, that is $\lambda_{RZA}\sim\mathcal{O}(\rho_{ZRA})$. Different MSD estimation performance curves are depicted as different parameters ρ_{ZRA}. In different SNR regimes, RZA-LMS can achieve the approximate optimal sparse channel estimation whose optimal regularization parameter is chosen as $\rho_{RZA} = 3$.

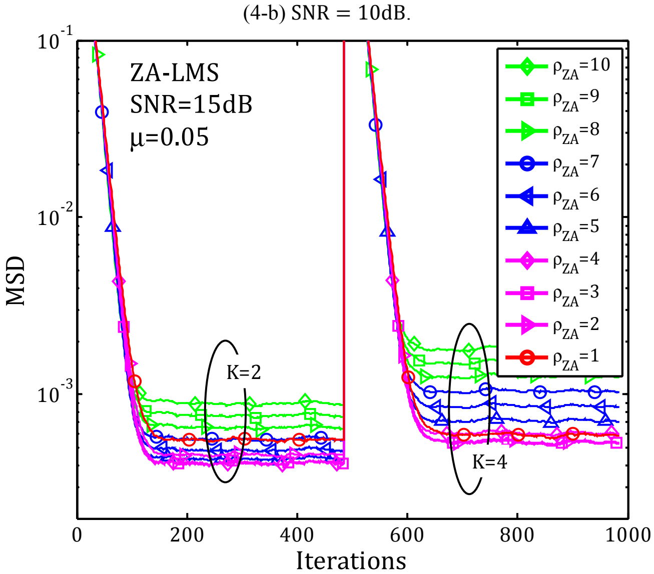
978-1-4673-6050-0/13/$31.00 ©2013 IEEE



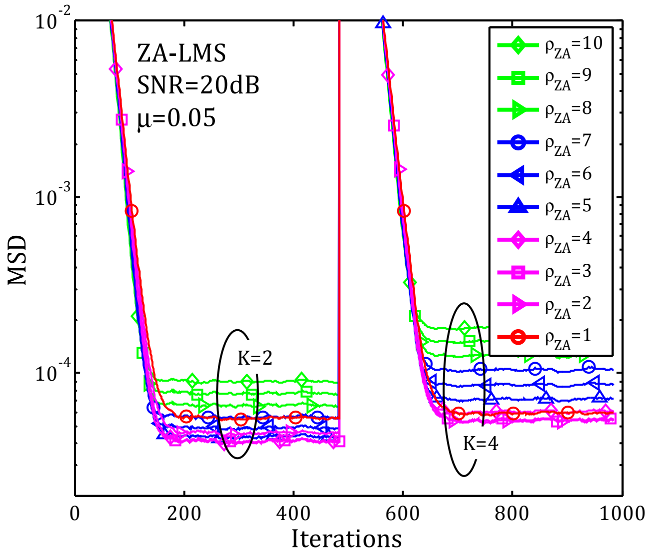
(4-a) SNR = 5dB.



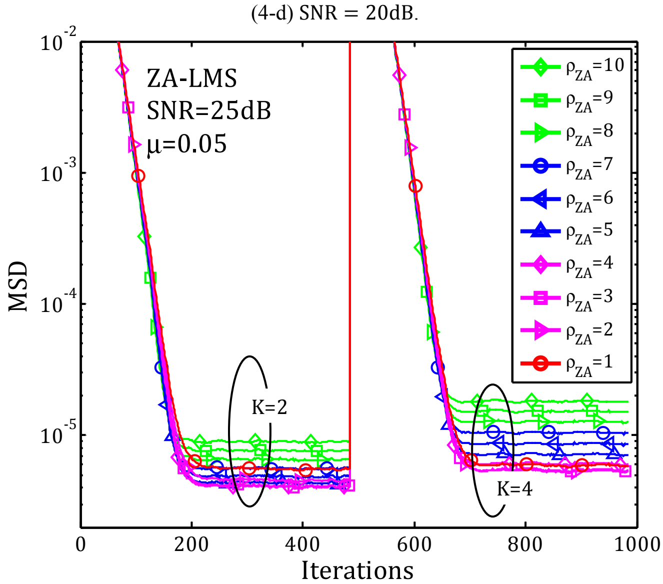
(4-b) SNR = 10dB.



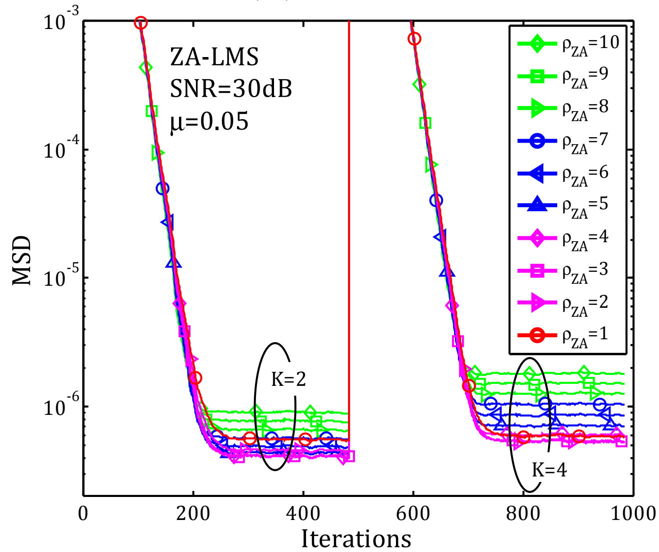
(4-c) SNR = 15dB.



(4-d) SNR = 20dB.

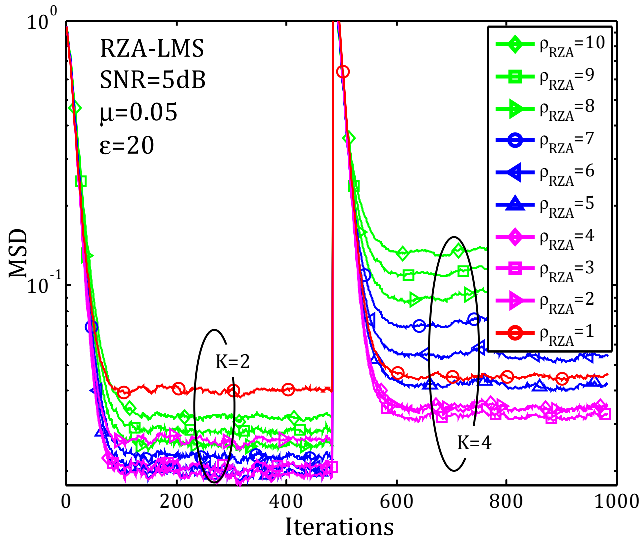


(4-e) SNR = 25dB.

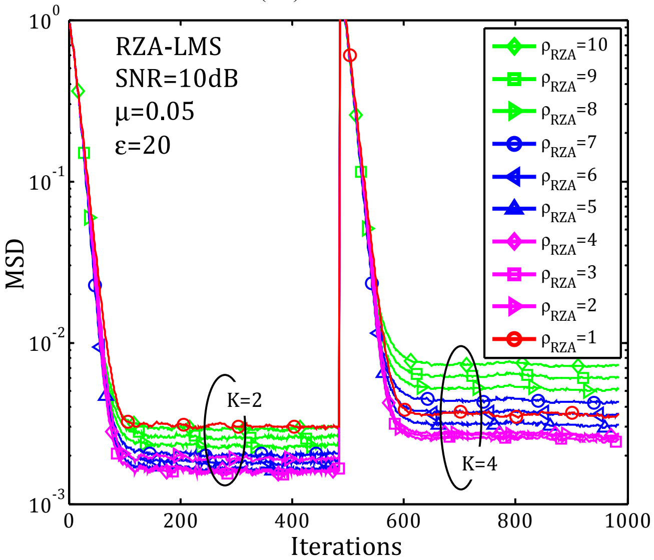


(4-f) SNR = 30dB.

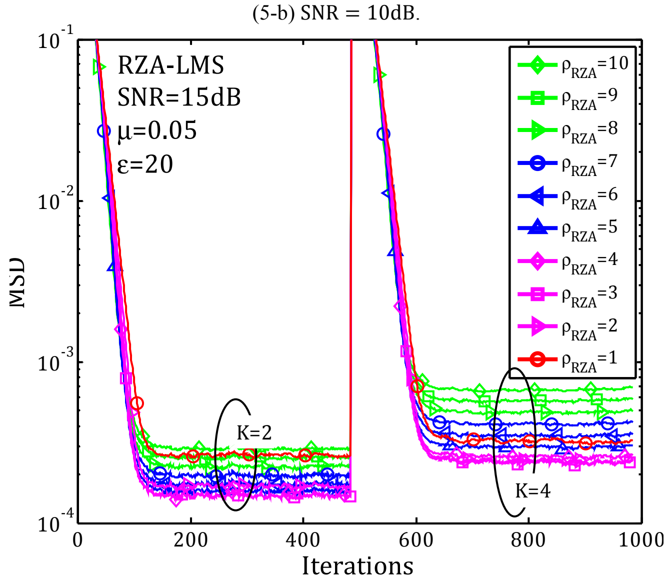
Fig. 4. MSD of ZA-LMS versus different regularization parameters.



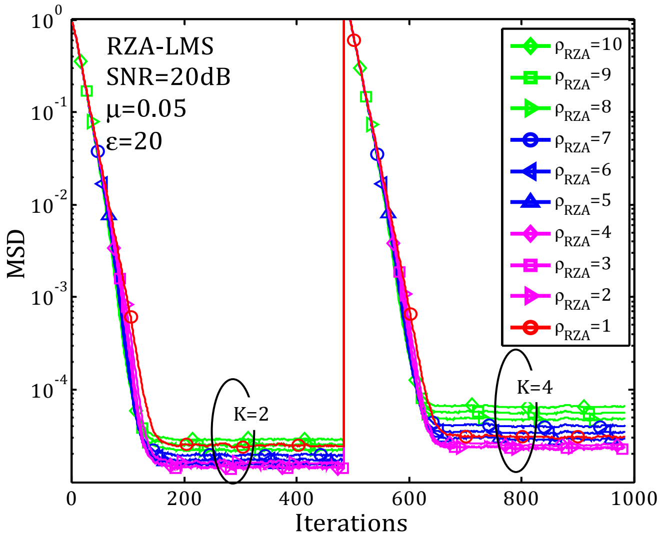
(5-a) SNR = 5dB.



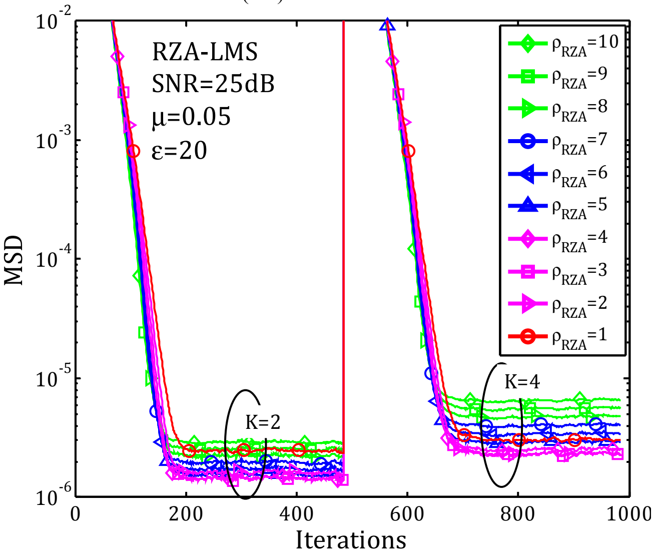
(5-b) SNR = 10dB.



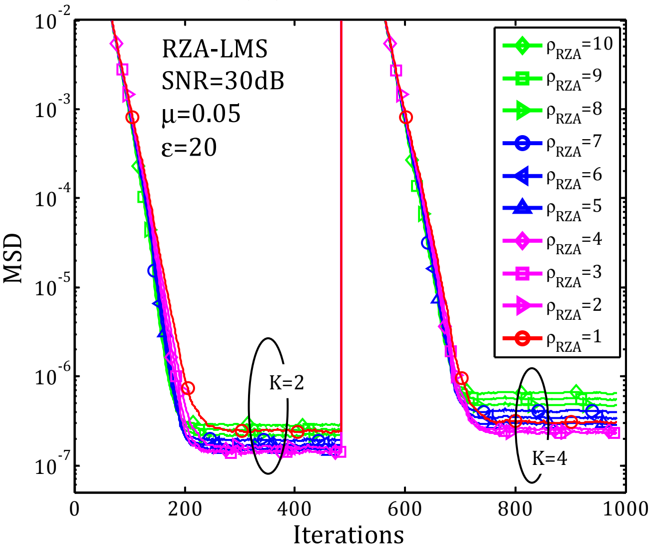
(5-c) SNR = 15dB.



(5-d) SNR = 20dB.



(5-e) SNR = 25dB.



(5-f) SNR = 30dB.

Fig. 5. MSD of RZA-LMS versus different regularization parameters.

V. Conclusion

By using ℓ_1-norm sparse constraint function, both ZA-LMS and RZA-LMS have been proposed for applying in sparse multipath channel estimation. We explained the relationship between LASSO and ℓ_1-norm based sparse LMS algorithms, i.e., ZA-LMS and RZA-LMS. Since the proposed methods neglect optimal regularization parameter selection. In this paper, we investigated regularization selection method for sparse LMS methods, i.e., ZA-LMS and RZA-LMS. Computer simulations were given to show the effectiveness of our propose method.

Acknowledment

This work was supported in part by the Japan Society for the Promotion of Science (JSPS) postdoctoral fellowship and the National Natural Science Foundation of China under Grant 61261048.

References

[1] D. Raychaudhuri and N. B. Mandayam, "Frontiers of wireless and mobile communications," *Proceedings of the IEEE*, vol. 100, no. 4, pp. 824–840, Apr. 2012.

[2] F. Adachi and E. Kudoh, "New direction of broadband wireless technology," *Wireless Communications and Mobile Computing*, vol. 7, no. 8, pp. 969–983, May 2007.

[3] F. Adachi, D. Grag, S. Takaoka, and K. Takeda, "Broadband CDMA techniques," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 8–18, Apr. 2005.

[4] N. Czink, X. Yin, H. OZcelik, M. Herdin, E. Bonek, and B. Fleury, "Cluster characteristics in a MIMO indoor propagation environment," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1465–1475, Apr. 2007.

[5] B. Widrow and D. Stearns, *Adaptive signal processing*, no. 4. New Jersey: Prentice Hall, 1985.

[6] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal frequency information incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[7] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[8] G. Taubock, F. Hlawatsch, D. Eiwen, and H. Rauhut, "Compressive estimation of doubly selective channels in multicarrier systems: Leakage effects and sparsity-enhancing processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 255–271, Apr. 2010.

[9] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.

[10] G. Gui, W. Peng, and F. Adachi, "High-resolution compressive channel estimation for broadband wireless communication systems," *International Journal of Communication Systems (WILEY)*, vol. 2012, pp. 1–12, Dec. 2012, doi: 10.1002/dac.2483.

[11] G. Gui, W. Peng, and F. Adachi, "Sub-Nyquist rate ADC sampling-based compressive channel estimation," *Wireless Communication and Mobile Computing,* pp. 1–10, *Apr. 2013, DOI:10.1002/wcm2372*.

[12] N. Wang, G. Gui, Z. Zhang, and T. Tang, "A novel sparse channel estimation method for multipath MIMO-OFDM systems," in *IEEE 74th Vehicular Technology Conference (VTC2011-Fall), San Francisco, California, USA*, 2011, pp. 1–5.

[13] G. Gui, A. Mehbodniya, and F. Adachi, "Bayesian sparse channel estimation and data detection for OFDM communication Systems," in *IEEE 78th Vehicular Technology Conference (VTC-Fall), 2-5 Sept. 2013, Las Vegas, USA*, 2013, pp. 1–5.

[14] E. J. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 1, no. 346, pp. 589–592, May 2008.

[15] R. Tibshirani, "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society (B)*, vol. 58, no. 1, pp. 267–288, 1996.

[16] Y. Chen, Y. Gu, and A. O. Hero III, "Sparse LMS for system identification," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Taipei, Taiwan, 19-24 April 2009, no. 3, pp. 3125–3128.

[17] G. Gui, W. Peng, and F. Adachi, "Improved adaptive sparse channel estimation based on the least mean square algorithm," in *IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, April 7-10*, 2013, pp. 3130–3134.

[18] G. Gui and F. Adachi, "Improved adaptive sparse channel estimation using least mean square algorithm," *EURASIP Journal on Wireless Communications and Networking, revised*, 2013.

[19] G. Gui and F. Adachi, "Adaptive Sparse Channel Estimation for Time-Variant MIMO-OFDM Systems," in *The 9th International Wireless Communications & Mobile Computing Conference (IWCMC), July 1 -5, 2013, Cagliari, Italy*, pp. 1–6.

[20] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l1 minimization," *Journal of Fourier Analysis Applications*, vol. 14, no. 5–6, pp. 877–905, Oct. 2008.