

RESEARCH

Open Access

# Improved least mean square algorithm with application to adaptive sparse channel estimation

Guan Gui\* and Fumiyuki Adachi

## Abstract

Least mean square (LMS)-based adaptive algorithms have attracted much attention due to their low computational complexity and reliable recovery capability. To exploit the channel sparsity, LMS-based adaptive sparse channel estimation methods have been proposed based on different sparse penalties, such as  $\ell_1$ -norm LMS or zero-attracting LMS (ZA-LMS), reweighted ZA-LMS, and  $\ell_p$ -norm LMS. However, the aforementioned methods cannot fully exploit channel sparse structure information. To fully take advantage of channel sparsity, in this paper, an improved sparse channel estimation method using  $\ell_0$ -norm LMS algorithm is proposed. The LMS-type sparse channel estimation methods have a common drawback of sensitivity to the scaling of random training signal. Thus, it is very hard to choose a proper learning rate to achieve a robust estimation performance. To solve this problem, we propose several improved adaptive sparse channel estimation methods using normalized LMS algorithm with different *sparse penalties*, which normalizes the power of input signal. Furthermore, Cramer-Rao lower bound of the proposed adaptive sparse channel estimator is derived based on prior information of channel taps' positions. Computer simulation results demonstrate the advantage of the proposed channel estimation methods in mean square error performance.

**Keywords:** Least mean square; Normalized LMS; Adaptive sparse channel estimation;  $\ell_p$ -Norm normalized least mean square;  $\ell_0$ -Norm normalized least mean square; Compressive sensing

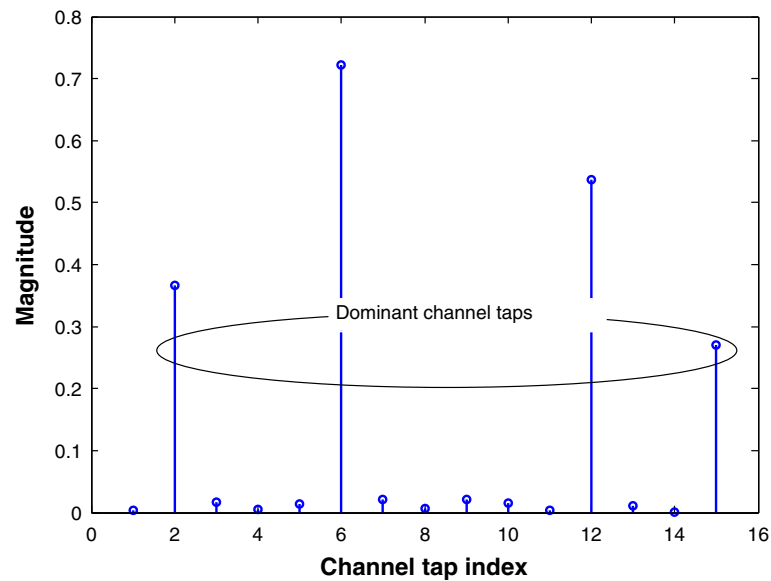
## 1 Introduction

The demand for high-speed data services has been increasing as emerging wireless devices are widely spreading. Various portable wireless devices, e.g., smartphones and laptops, have generated rapidly, giving rise to massive data traffic [1]. Broadband transmission is an indispensable technique in the next-generation wireless communication systems [2,3]. The broadband channel is described by a sparse channel model in which multipath taps are widely separated in time, thereby creating a large delay spread [4-8]. In other words, most of channel coefficients are zero or close to zero, while only a few channel coefficients are *dominant* (large value). A typical example of sparse multipath channel is shown in Figure 1, where the number of dominant taps is four while the length is 16.

Traditional least mean square (LMS) is one of the most popular algorithms for adaptive system identification [9], e.g., channel estimation. LMS-based adaptive channel

estimation can be easily implemented due to its low computational complexity. In current broadband wireless communication systems, channel impulse response in time domain is often described by a sparse channel model, supported by a few large coefficients. The LMS-based adaptive channel estimation method never takes advantage of the channel sparse structure although its mean square error (MSE) lower bound has a direct relationship with finite impulse response (FIR) channel length. To improve the estimation performance, recently, many algorithms have been proposed to take advantage of the sparse nature of the channel. For example, based on the recent theory of compressive sensing (CS) [10,11], various sparse channel estimation methods have been proposed in [12-14]. Some of these methods are known to achieve robust estimation, e.g., sparse channel estimation using least-absolute shrinkage and selection operator [15]. However, these kinds of sparse methods have two potential disadvantages: one is that computational complexity may be very high, especially for tracking fast time-variant channels; the other is that training signal matrices for these CS-based sparse channel

\* Correspondence: [gui@mobile.ecei.tohoku.ac.jp](mailto:gui@mobile.ecei.tohoku.ac.jp)  
Department of Communications Engineering Graduate School of Engineering, Tohoku University, Sendai 980-8579, Japan



**Figure 1** A typical example of sparse multipath channel.

estimation methods are required to satisfy the restricted isometry property [16]. It was well known that designing these training matrices is a non-deterministic polynomial-time (NP)-hard problem [17].

In order to avoid the two detrimental problems of sparse channel estimation, a variation of the LMS algorithm with  $\ell_1$ -norm sparse constraint has been proposed in [18]. The  $\ell_1$ -norm sparse penalty is incorporated into the cost function of the conventional LMS algorithm, which results in the LMS update with a zero attractor, namely *zero-attracting LMS* (ZA-LMS) and *reweighted ZA-LMS* (RZA-LMS) [18] which is motivated by the reweighted  $\ell_1$ -norm minimization recovery algorithm [19]. To further improve the estimation performance, an adaptive sparse channel estimation method using  $\ell_p$ -norm LMS (LP-LMS) algorithm has been proposed [20]. However, there still exists a performance gap between LP-LMS and optimal sparse channel estimation. It is worth mentioning that optimal channel estimation is often denoted by the sparse lower bound which is derived in Theorem 4 in Section 3.

Due to the fact that solving the LP-LMS algorithm is a non-convex problem, the algorithm cannot obtain an optimal adaptive sparse solution [10,11]. Hence, a computationally efficient algorithm is required to obtain a more accurate adaptive sparse channel estimation.

According to the CS theory [10,11], solving the  $\ell_0$ -norm sparse penalty problem can obtain an optimal sparse solution. In other words,  $\ell_0$ -norm sparse penalty on LMS algorithm is a good candidate to achieve more accurate channel estimations. These backgrounds motivate us to use the optimal  $\ell_0$ -norm

sparse penalty on LMS in order to improve the estimation performance.

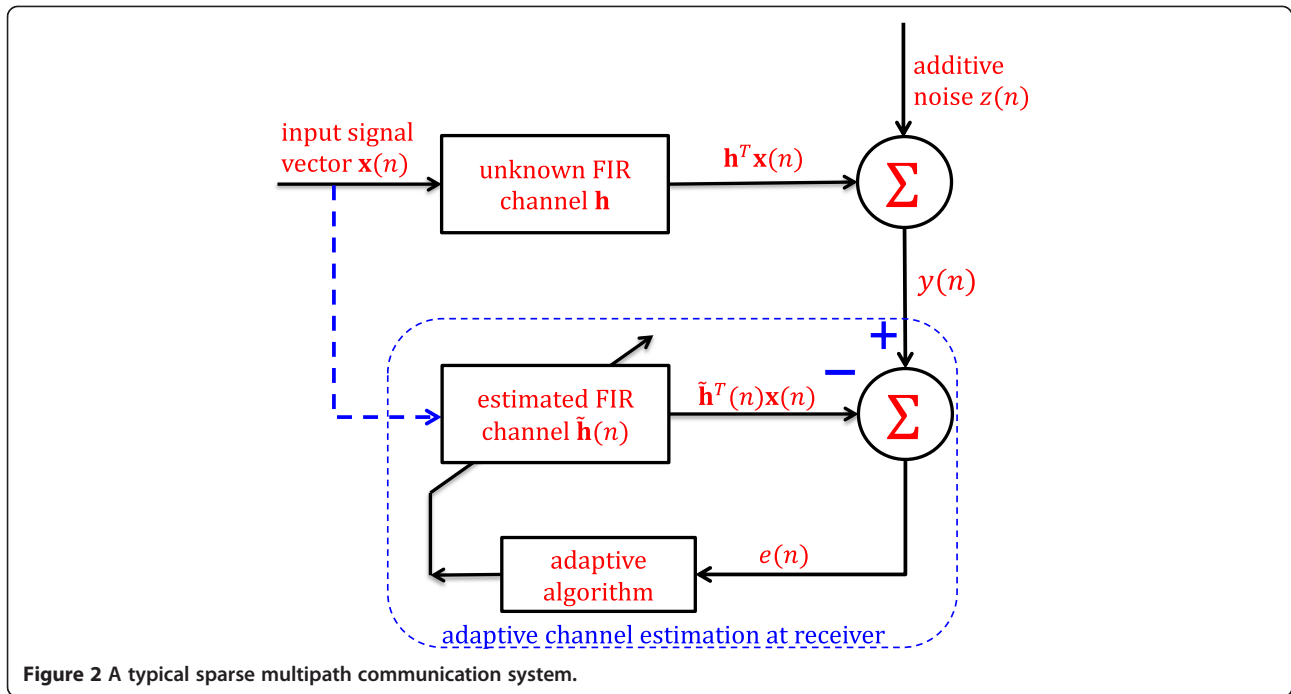
In addition, since sparse LMS-based channel estimation methods have a common drawback of sensitivity to scaling of random training signal, it is very hard to choose a proper learning rate to achieve a robust estimation performance [21]. To solve this problem, we propose several improved adaptive sparse channel estimation methods using *normalized LMS* (NLMS) algorithm, which normalizes the power of input signal, with different sparse penalties, i.e.,  $\ell_p$ -norm ( $0 \leq p \leq 1$ ).

The contributions of this paper are described below. Firstly, we propose an improved adaptive sparse channel estimation method using  $\ell_0$ -norm least square error algorithm, termed as L0-LMS [22]. Secondly, based on algorithms in [18,20], we propose four kinds of improved adaptive sparse channel estimation methods using sparse NLMS algorithms. Thirdly, Cramer-Rao lower bound (CRLB) of the proposed adaptive sparse channel estimator is derived based on prior information of known positions of non-zero taps. Lastly, various simulation results are given to confirm the effectiveness of our proposed methods.

Section 2 introduces the system model and problem formulation. Section 3 discusses various adaptive sparse channel estimation methods using different LMS-based algorithms. In Section 4, computer simulation results for the MSE performance are presented to confirm the effectiveness of sparsity-aware modifications of LMS. Concluding remarks are presented in Section 5.

## 2 System model and problem formulation

Consider a sparse multipath communication system, as shown in Figure 2.



**Figure 2** A typical sparse multipath communication system.

The input signal  $\mathbf{x}(t)$  and ideal output signal  $y(t)$  are related by

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

where  $\mathbf{h} = [h_0, h_1, \dots, h_{N-1}]^T$  is  $N$ -length sparse channel vector which is supported by  $K$  ( $K \ll N$ ) dominant taps,  $\mathbf{x}(t) = [x(t), x(t-1), \dots, x(t-N+1)]^T$  is an  $N$ -length input signal vector of  $x(t)$  and  $z(t)$  is an additive noise at time  $t$ . The objective of the LMS-type adaptive filter is to estimate the unknown sparse channel  $\mathbf{h}$  using  $\mathbf{x}(t)$  and  $y(t)$ . According to Equation 1, the  $n$ th channel estimation error  $e(n)$  is easily written as

$$e(n) = y(n) - \tilde{\mathbf{h}}^T(n) \mathbf{x}(n) = y(n) - \tilde{\mathbf{h}}^T(n) \mathbf{x}(n), \quad (2)$$

at time  $t$ , where  $\tilde{\mathbf{h}}(n)$  is the LMS adaptive channel estimator. It was worth noting that  $\mathbf{x}(t)$  and  $y(t)$  are invariant as iterative times. Based on Equation 2, standard LMS cost function can be written as

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

Hence, the updated equation of LMS adaptive channel estimation is derived as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) - \mu \frac{\partial L(n)}{\partial \tilde{\mathbf{h}}(n)} = \tilde{\mathbf{h}}(n) + \mu e(n) \mathbf{x}(n), \quad (4)$$

where  $\mu$  is the step size of gradient descend. For later use, we define a parameter  $\gamma_{\max}$  which is the maximum eigenvalue of the covariance matrix  $\mathbf{R} = E\{\mathbf{x}(t) \mathbf{x}^T(t)\}$  of

input signal vector  $\mathbf{x}(t)$ . Here, we introduce Theorem 1 in relation to the step size  $\mu$ . The detailed derivation can also be found in [21].

**Theorem 1** *The necessary condition of reliable LMS adaptive channel estimation is  $0 < \mu < 2/\gamma_{\max}$ .*

*Proof* At first, the  $(n+1)$ th update coefficient vector  $\tilde{\mathbf{h}}(n+1)$  is written as

$$\begin{aligned} \tilde{\mathbf{h}}(n+1) &= \tilde{\mathbf{h}}(n) + \mu \mathbf{x}(n) e(n) \\ &= \tilde{\mathbf{h}}(n) + \mu \mathbf{x}(n) (y(n) - \tilde{\mathbf{h}}^T(n) \mathbf{x}(n)). \end{aligned} \quad (5)$$

Assume that  $\mathbf{h}$  is a real FIR channel vector; subtracting  $\mathbf{h}$  from both sides in Equation 5, we can rewrite it as

$$\tilde{\mathbf{h}}(n+1) - \mathbf{h} = \tilde{\mathbf{h}}(n) - \mathbf{h} + \mu \mathbf{x}(n) (y(n) - \tilde{\mathbf{h}}^T(n) \mathbf{x}(n)). \quad (6)$$

By defining  $\mathbf{v}(n) = \tilde{\mathbf{h}}(n) - \mathbf{h}$  and according to Equation 1, Equation 6 can be rewritten as

$$\begin{aligned} \mathbf{v}(n+1) &= \mathbf{v}(n) + \mu \mathbf{x}(n) (y(n) - \tilde{\mathbf{h}}^T(n) \mathbf{x}(n)) \\ &\quad + \mu \mathbf{x}(n) (\mathbf{x}^T(n) \mathbf{h} - \mathbf{x}^T(n) \tilde{\mathbf{h}}(n)) \\ &= \mathbf{v}(n) + \mu \mathbf{x}(n) z(n) - \mu \mathbf{x}(n) \mathbf{x}^T(n) \mathbf{v}(n) \\ &= (\mathbf{I} - \mu \mathbf{x}(n) \mathbf{x}^T(n)) \mathbf{v}(n) + \mu \mathbf{x}(n) z(n). \end{aligned} \quad (7)$$

Assume that input signal vector  $\mathbf{x}(t)$  and estimated FIR channel vector  $\tilde{\mathbf{h}}(n)$  are statistically independent

from the other, then  $\mathbf{x}(t)$  and  $\mathbf{z}(t)$  are also independent, i.e.,  $E\{\mathbf{x}(t)\mathbf{z}(t)\} = 0$ . Taking the expectation for both sides in Equation 7, we can obtain

$$\begin{aligned} E\{\mathbf{v}(n+1)\} &= E\{(\mathbf{I}-\mu\mathbf{x}(t)\mathbf{x}^T(t))\mathbf{v}(n)\} + E\{\mu\mathbf{x}(t)\mathbf{z}(t)\} \\ &= (\mathbf{I}-\mu E\{\mathbf{x}(t)\mathbf{x}^T(t)\})E\{\mathbf{v}(n)\} + E\{\mu\mathbf{x}(t)\mathbf{z}(t)\} \\ &= (\mathbf{I}-\mu\mathbf{R})E\{\mathbf{v}(n)\}, \end{aligned} \quad (8)$$

The necessary condition of the reliable LMS adaptive channel estimation is given by

$$|1-\mu\lambda_{\max}| < 1, \quad (9)$$

where  $\lambda_{\max}$  denotes the maximum eigenvalue of  $\mathbf{R}$ . Hence, to guarantee fast convergence speed and reliable updating of LMS in Equation 4, step size  $\mu$  should be satisfied:

$$0 < \mu < 2/\lambda_{\max}. \quad (10)$$

□

### 3 (N)LMS-based adaptive sparse channel estimation

From Equation 4, we can find that the LMS-based channel estimation method never takes advantage of the sparse structure in channel vector  $\mathbf{h}$ . To get a better understanding, the LMS-based channel estimation methods can be expressed as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \text{Adaptive error update}. \quad (11)$$

Unlike the conventional LMS method in Equation 11, sparse LMS algorithms exploit channel sparsity by introducing several  $\ell_p$ -norm penalties to their cost functions with  $0 \leq p \leq 1$ . The LMS-based adaptive sparse channel estimation methods can be written as

$$\tilde{\mathbf{h}}(n+1) = \underbrace{\tilde{\mathbf{h}}(n) + \text{Adaptive error update}}_{\text{LMS}} + \underbrace{\text{Sparse penalty}}_{\text{sparse LMS}}. \quad (12)$$

Equation 12 motivated us to introduce different sparse penalties in order to take advantage of the sparse structure as prior information. Here, if we analogize the updated equation in Equation 12 to the CS-based sparse channel estimation [10,11], one can find that more accurate sparse channel approximation is adopted, and then better estimation accuracy could be obtained and vice versa. The conventional sparse penalties are  $\ell_p$ -norm ( $0 < p \leq 1$ ) and  $\ell_0$ -norm, respectively. Since  $\ell_0$ -norm penalty on sparse signal recovery is a well-known NP-hard problem, only the  $\ell_p$ -norm ( $0 < p \leq 1$ )-based sparse LMS approaches have been proposed for adaptive sparse channel estimation. Compared to conventional two sparse LMS algorithms (ZA-LMS and RZA-LMS), LP-LMS can achieve a better

estimation performance. However, there still exists a performance gap between the LP-LMS-based channel estimator and the optimal one. As the development of mathematics continues, a more accurate sparse approximate algorithm to  $\ell_0$ -norm LMS (L0-LMS) is proposed in [22] and analyzed in [23]. However, they never considered any application on sparse channel estimation. In this paper, the L0-LMS algorithm is applied in adaptive sparse channel estimation to improve the estimation performance.

It is easily found that exploitation of more accurate sparse structure information can obtain a better estimation performance. In the following, we investigate sparse LMS-based adaptive sparse channel estimation methods using different sparse penalties.

#### 3.1 LMS-based adaptive sparse channel estimation

The following are the LMS-based adaptive sparse channel estimation methods:

- *ZA-LMS*. To exploit the channel sparsity in time domain, the cost function of ZA-LMS [18] is given by

$$L_{ZA}(n) = \frac{1}{2}e^2(n) + \lambda_{ZA}\|\tilde{\mathbf{h}}(n)\|_1, \quad (13)$$

where  $\lambda_{ZA}$  is a regularization parameter to balance the estimation error and sparse penalty of  $\tilde{\mathbf{h}}(n)$ . The corresponding updated equation of ZA-LMS is

$$\begin{aligned} \tilde{\mathbf{h}}(n+1) &= \tilde{\mathbf{h}}(n) - \mu \frac{\partial L_{ZA}(n)}{\partial \tilde{\mathbf{h}}(n)} \\ &= \tilde{\mathbf{h}}(n) + \mu e(n)\mathbf{x}(t) - \rho_{ZA} \text{sgn}(\tilde{\mathbf{h}}(n)), \end{aligned} \quad (14)$$

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

$$\text{sgn}(\mathbf{h}) = \begin{cases} 1, & h > 0 \\ 0, & h = 0 \\ -1, & h < 0 \end{cases}.$$

From the updated equation in Equation 14, the function of its second term is compressing small channel coefficients as zero in high probability. That is to say, most of the small channel coefficients can be simply replaced by zeros, which speeds up the convergence of this algorithm.

- *RZA-LMS*. ZA-LMS cannot distinguish between zero taps and non-zero taps as it gives the same penalty to all the taps which are often forced to be zero with the same probability; therefore, its performance will degrade in less sparse systems. Motivated by the reweighted  $\ell_1$ -norm minimization recovery algorithm [19], Chen et al. have proposed a heuristic approach to

reinforce the zero attractor which was termed as the RZA-LMS [18]. The cost function of RZA-LMS is given by

$$L_{RZA}(n) = \frac{1}{2}e^2(n) + \lambda_{RZA} \sum_{i=1}^N \log(1 + \varepsilon_{RZA}|h_i(n)|), \quad (15)$$

where  $\lambda_{RZA} > 0$  is the regularization parameter and  $\varepsilon_{RZA} > 0$  is the positive threshold. In computer simulation, the threshold is set as  $\varepsilon_{RZA} = 20$  which is also suggested in [18]. The  $i$ th channel coefficient  $\tilde{h}_i(n)$  is then updated as

$$\begin{aligned} \tilde{h}_i(n+1) &= \tilde{h}_i(n) - \mu \frac{\partial L_{RZA}(n)}{\partial \tilde{h}_i(n)} \\ &= \tilde{h}_i(n) + \mu e(n) \mathbf{x}(t-i) - \rho_{RZA} \frac{\text{sgn}(\tilde{h}_i(n))}{1 + \varepsilon_{RZA}|\tilde{h}_i(n)|}, \end{aligned} \quad (16)$$

where  $\rho_{RZA} = \mu \lambda_{RZA} \varepsilon_{RZA}$ . Equation 16 can be expressed in the vector form as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mu e(n) \mathbf{x}(t) - \rho_{RZA} \frac{\text{sgn}(\tilde{\mathbf{h}}(n))}{1 + \varepsilon_{RZA}|\tilde{\mathbf{h}}(n)|}.$$

Please note that the second term in Equation 16 attracts the channel coefficients  $\tilde{h}_i(n)$ ,  $i = 1, 2, \dots, N$  whose magnitudes are comparable to  $1/\varepsilon_{RZA}$  to zeros.

- *LP-LMS*. Following the above ideas in [18], LP-LMS-based adaptive sparse channel estimation method has been proposed in [20]. The cost function of LP-LMS is given by

$$L_{LP}(n) = \frac{1}{2}e^2(n) + \lambda_{LP} \|\tilde{\mathbf{h}}(n)\|_p, \quad (17)$$

where  $\lambda_{LP} > 0$  is a regularization parameter. The corresponding updated equation of LP-LMS is given as

$$\begin{aligned} \tilde{\mathbf{h}}(n+1) &= \tilde{\mathbf{h}}(n) - \mu \frac{\partial L_{LP}(n)}{\partial \tilde{\mathbf{h}}(n)} \\ &= \tilde{\mathbf{h}}(n) + \mu e(n) \mathbf{x}(t) - \rho_{LP} \frac{\|\tilde{\mathbf{h}}(n)\|_p^{1-p} \text{sgn}\{\tilde{\mathbf{h}}(n)\}}{\varepsilon_{LP} + \|\tilde{\mathbf{h}}(n)\|_p^{(1-p)}}, \end{aligned} \quad (18)$$

where  $\rho_{LP} = \mu \lambda_{LP}$  and  $\varepsilon_{LP} > 0$  is a small positive parameter.

- *L0-LMS (proposed)*. Consider  $\ell_0$ -norm penalty on LMS cost function to produce a sparse channel estimator as this penalty term forces the channel tap values of  $\tilde{\mathbf{h}}(n)$  to approach zero. The cost function of L0-LMS is given by

$$L_{L0}(n) = \frac{1}{2}e^2(n) + \lambda_{L0} \|\tilde{\mathbf{h}}(n)\|_0, \quad (19)$$

where  $\lambda_{L0} > 0$  is a regularization parameter and  $\|\tilde{\mathbf{h}}(n)\|_0$  denotes  $\ell_0$ -norm sparse penalty function which counts the number of non-zero channel taps of  $\tilde{\mathbf{h}}(n)$ . Since solving the  $\ell_0$ -norm minimization is an NP-hard problem [17], to reduce computational complexity, we replace it with an approximate continuous function:

$$\|\tilde{\mathbf{h}}(n)\|_0 \approx \sum_{i=0}^{N-1} \left(1 - e^{-\beta|\tilde{h}_i(n)|}\right). \quad (20)$$

The cost function in Equation 19 can then be rewritten as

$$L_{L0}(n) = \frac{1}{2}e^2(n) + \lambda_{L0} \sum_{i=0}^{N-1} \left(1 - e^{-\beta|\tilde{h}_i(n)|}\right). \quad (21)$$

The first-order Taylor series expansion of exponential function  $e^{-\beta|\tilde{h}_i(n)|}$  is given as

$$e^{-\beta|\tilde{h}_i(n)|} \approx \begin{cases} 1 - \beta|\tilde{h}_i(n)|, & \text{when } |\tilde{h}_i(n)| \leq \frac{1}{\beta}. \\ 0, & \text{otherwise} \end{cases} \quad (22)$$

The updated equation of L0-LMS-based adaptive sparse channel estimation can then be derived as

$$\begin{aligned} \tilde{\mathbf{h}}(n+1) &= \tilde{\mathbf{h}}(n) \\ &+ \mu e(n) \mathbf{x}(t) - \rho_{L0} \beta \text{sgn}(\tilde{\mathbf{h}}(n)) e^{-\beta|\tilde{\mathbf{h}}(n)|}, \end{aligned} \quad (23)$$

where  $\rho_{L0} = \mu \lambda_{L0}$ . Unfortunately, the exponential function in Equation 23 will also cause high computational complexity. To further reduce the complexity, an approximation function  $J(\tilde{\mathbf{h}}(n))$  is also proposed to the updated Equation 23. Finally, the updated equation of L0-LMS-based adaptive sparse channel estimation can be derived as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mu e(n) \mathbf{x}(t) - \rho_{L0} J(\tilde{\mathbf{h}}(n)), \quad (24)$$

where  $\rho_{L0} = \mu \lambda_{L0}$  and  $J(\tilde{\mathbf{h}}(n))$  is defined as

$$J(\tilde{\mathbf{h}}(n)) = \begin{cases} 2\beta^2 \tilde{h}_i(n) - 2\beta \cdot \text{sgn}(\tilde{h}_i(n)) & \text{when } |\tilde{h}_i(n)| \leq \frac{1}{\beta} \\ 0, & \text{otherwise,} \end{cases} \quad (25)$$

for all  $i \in \{1, 2, \dots, N\}$ .

### 3.2 Improved adaptive sparse channel estimation methods

The common drawback of the above sparse LMS-based algorithms is that they are vulnerable to probabilistic



scaling of their training signal  $\mathbf{x}(t)$ . In other words, LMS-based algorithms are sensitive to signal scaling [21]. Hence, it is very hard to choose a proper step size  $\mu$  to guarantee stability of these sparse LMS-based algorithms even if the step size satisfies the necessary condition in Equation 10.

Let us reconsider the updated equation of LMS in Equation 4. Assuming that the  $n$ th adaptive channel estimator  $\tilde{\mathbf{h}}(n+1)$  is the optimal solution, the relationship between the  $(n+1)$ th channel estimator  $\tilde{\mathbf{h}}(n+1)$  and input signal  $\mathbf{x}(t)$  is given as

$$\tilde{\mathbf{h}}^T(n+1)\mathbf{x}(t) = y(t), \quad (26)$$

where  $y(t)$  is assumed to be ideal received signal at the receiver. To solve a convex optimization problem in Equation 26, the cost function can be constructed as [21]

$$C(n) = (\tilde{\mathbf{h}}(n+1) - \tilde{\mathbf{h}}(n))^T (\tilde{\mathbf{h}}(n+1) - \tilde{\mathbf{h}}(n)) + \xi(y(t) - \tilde{\mathbf{h}}^T(n+1)\mathbf{x}(t)), \quad (27)$$

where  $\xi$  is the unknown real-value Lagrange multiplier [21]. The optimal channel estimator at the  $(n+1)$ th update can be found by letting the first derivative of

$$C(n) = 0. \quad (28)$$

Hence, it can be derived as

$$\frac{\partial C(n)}{\partial \tilde{\mathbf{h}}(n+1)} = 2(\tilde{\mathbf{h}}(n+1) - \tilde{\mathbf{h}}(n)) - \xi\mathbf{x}(t) = 0. \quad (29)$$

The  $(n+1)$ th optimal channel estimator is given from Equation 29 as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \frac{1}{2}\xi\mathbf{x}(t). \quad (30)$$

By substituting Equation 30 into Equation 26, we obtain

$$\xi\mathbf{x}^T(t)\mathbf{x}(t) = 2(y(t) - \tilde{\mathbf{h}}^T(n)\mathbf{x}(t)) = 2e(n), \quad (31)$$

where  $e(n) = y(t) - \tilde{\mathbf{h}}^T(n)\mathbf{x}(t)$  (see Equation 2) and the unknown parameter  $\xi$  is given by

$$\xi = \frac{2e(n)}{\mathbf{x}^T(t)\mathbf{x}(t)}. \quad (32)$$

By substituting it to Equation 30, the updated equation of NLMS is written as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mu_1 \frac{e(n)\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)}, \quad (33)$$

where  $\mu_1$  is the gradient step size which controls the adaptive convergence speed of NLMS algorithm. Based on the updated Equation 33, for better understanding, NLMS-based sparse adaptive updated equation can be generalized as

$$\tilde{\mathbf{h}}(n+1) = \underbrace{\tilde{\mathbf{h}}(n)}_{\text{NLMS}} + \underbrace{\text{Normalized adaptive error update}}_{\text{sparse NLMS}} + \text{Sparse penalty}, \quad (34)$$

where normalized adaptive update term is  $\mu_1 e(n)\mathbf{x}(t)/\mathbf{x}^T(t)\mathbf{x}(t)$  which replaces the adaptive update  $\mu e(n)\mathbf{x}(t)$  in Equation 4. The advantage of NLMS-based adaptive sparse channel estimation is that it can mitigate the scaling interference of training signal due to the fact that NLMS-based methods estimate the sparse channel by normalizing with the power of training signal  $\mathbf{x}(t)$ . To ensure the stability of the NLMS-based algorithms, the necessary condition of step size  $\mu_1$  is derived briefly. The detail derivation can also be found in [21].

**Theorem 2** *The necessary condition of reliable NLMS adaptive channel estimation is*

$$0 < \mu_1 < \frac{E\{e(n)(\mathbf{h} - \tilde{\mathbf{h}}(n))^T \mathbf{x}(t) / \mathbf{x}^T(t)\mathbf{x}(t)\}}{E\{e^2(n) / \mathbf{x}^T(t)\mathbf{x}(t)\}}. \quad (35)$$

*Proof* Since the NLMS-based algorithms share the same gradient step size to ensure their stability, for simplicity, studying the NLMS for a general case. The updated equation of NLMS is given by

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mu_1 e(n) \frac{\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)}, \quad (36)$$

where  $\mu_1$  denotes the step size of NLMS-type algorithms. Denoting the channel estimation error vector as  $\mathbf{u}(n) = \mathbf{h} - \tilde{\mathbf{h}}(n)$ ,  $(n+1)$ th update error  $\mathbf{u}(n+1)$  can be written as

$$\mathbf{u}(n+1) = \mathbf{u}(n) - \mu_1 e(n) \frac{\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)}. \quad (37)$$

Obviously, the  $(n+1)$ th update MSE  $E\{\mathbf{u}^2(n+1)\}$  can also be given by

$$E\{\mathbf{u}^2(n+1)\} = E\{\mathbf{u}^2(n)\} - 2\mu_1 E\left\{\frac{e(n)\mathbf{u}^T(n)\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)}\right\} + \mu_1^2 E\left\{\frac{e^2(n)}{\mathbf{x}^T(t)\mathbf{x}(t)}\right\}. \quad (38)$$

To ensure the stable updating of the NLMS-type algorithms, the necessary condition is satisfying

$$E\{\mathbf{u}^2(n+1)\} - E\{\mathbf{u}^2(n)\} = -2\mu_1 E\left\{\frac{e(n)(\mathbf{h}-\tilde{\mathbf{h}}(n))^T \mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)}\right\} + \mu_1^2 E\left\{\frac{e^2(n)}{\mathbf{x}^T(t)\mathbf{x}(t)}\right\} \leq 0. \quad (39)$$

Hence, the necessary condition of reliable adaptive sparse channel estimation is  $\mu_1$  satisfying the theorem in Equation 35.  $\square$

The following are the improved adaptive sparse channel estimation methods:

- *ZA-NLMS (proposed)*. According to the Equation 14, the updated equation of ZA-NLMS can be written as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mu_1 e(n) \frac{\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)} - \rho_{ZAN} \text{sgn}(\tilde{\mathbf{h}}(n)). \quad (40)$$

where  $\rho_{ZAN} = \mu_1 \gamma_{ZAN}$  and  $\gamma_{ZAN}$  is a regularization parameter for ZA-NLMS.

- *RZA-NLMS (proposed)*. According to Equation 16, the updated equation of RZA-NLMS can be written as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mu_1 e(n) \frac{\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)} - \rho_{RZAN} \frac{\text{sgn}(\tilde{\mathbf{h}}(n))}{1 + \varepsilon_{RZA} |\tilde{\mathbf{h}}(n)|}. \quad (41)$$

where  $\rho_{RZAN} = \mu_1 \lambda_{RZA} \varepsilon_{RZAN}$  and  $\gamma_{RZAN}$  is a regularization parameter for RZA-NLMS. The threshold is set as  $\varepsilon_{RZAN} = \varepsilon_{RZA} = 20$  which is also consistent with our previous research in [24-27].

- *LP-NLMS (proposed)*. According to the LP-LMS in Equation 18, the updated equation of LP-NLMS can be written as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mu_1 e(n) \frac{\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)} - \rho_{LPN} \frac{\|\tilde{\mathbf{h}}(n)\|_p^{1-p} \text{sgn}\{\tilde{\mathbf{h}}(n)\}}{\varepsilon_{LPN} + \|\tilde{\mathbf{h}}(n)\|^{(1-p)}}, \quad (42)$$

where  $\rho_{LPN} = \mu_1 \lambda_{LPN}/10$ ,  $\lambda_{LPN}$  is a regularization parameter, and  $\varepsilon_{LPN} > 0$  is a threshold parameter.

- *L0-NLMS (proposed)*. Based on updated the equation of L0-LMS algorithm in Equation 24, the updated equation of L0-NLMS algorithm can be directly written as

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mu_1 e(n) \frac{\mathbf{x}(t)}{\mathbf{x}^T(t)\mathbf{x}(t)} - \rho_{L0N} J(\tilde{\mathbf{h}}(n)), \quad (43)$$

where  $\rho_{L0N} = \mu_1 \lambda_{L0N}$  and  $\lambda_{L0N}$  is a regularization parameter. The sparse penalty function  $J(\tilde{\mathbf{h}}(n))$  has been defined as in (25).

### 3.3 Cramer-Rao lower bound

To decide the CRLB of the proposed channel estimator, Theorems 3 and 4 are derived as follows.

**Theorem 3** For an  $N$ -length channel vector  $\mathbf{h}$ , if  $\mu$  satisfies  $0 < \mu < 2/\lambda_{max}$ , then MSE lower bound of LMS adaptive channel estimator is  $B = \mu P_0 N / (2 - \mu \lambda_{min}) \sim \mathcal{O}(N)$ , where  $P_0$  is a parameter which denotes unit power of gradient noise and  $\lambda_{min}$  denotes the minimum eigenvalue of  $\mathbf{R}$ .

*Proof* Firstly, we define the estimation error at the  $(n+1)$ th iteration  $\mathbf{v}(n+1)$  as

$$\begin{aligned} \mathbf{v}(n+1) &= \tilde{\mathbf{h}}(n+1) - \mathbf{h} \\ &= \mathbf{v}(n) + \mu \mathbf{x}(t) e(n) \\ &= \mathbf{v}(n) - \frac{1}{2} \mu \Gamma(n), \end{aligned} \quad (44)$$

where  $\Gamma(n) = \partial L(n) / \partial \tilde{\mathbf{h}}(n) = -2\mathbf{x}(t)e(n)$  is a joint gradient error function which includes channel estimation error and noise plus interference error. To derive the lower bound of the channel estimator, two gradient errors should be separated. Hence, assuming  $\Gamma(n)$  can be split in two terms:  $\Gamma(n) = \tilde{\Gamma}(n) + 2\mathbf{w}(n)$  where  $\tilde{\Gamma}(n) = 2(\mathbf{R}\tilde{\mathbf{h}}(n) - \mathbf{p})$  denotes the gradient error and  $\mathbf{w}(n) = [w_0(n), w_1(n), \dots, w_{N-1}(n)]^T$  represents the gradient noise vector [21]. Obviously,  $E\{\mathbf{w}(n)\} = 0$  and

$$\begin{aligned} \mathbf{x}(t)e(n) &= -\frac{1}{2} \Gamma(n) \\ &= -\mathbf{R}(\tilde{\mathbf{h}}(n) - \mathbf{h}) - \mathbf{w}(n) \\ &= -(\mathbf{R}\tilde{\mathbf{h}}(n) - \mathbf{p}) - \mathbf{w}(n) \\ &= -\mathbf{R}(\tilde{\mathbf{h}}(n) - \mathbf{h}) - \mathbf{w}(n) \\ &= -\mathbf{R}\mathbf{v}(n) - \mathbf{w}(n), \end{aligned} \quad (45)$$

where  $\mathbf{p} = \mathbf{R}\mathbf{h}$ . Then, we rewrite  $\mathbf{v}(n+1)$  in Equation 44 as

$$\begin{aligned}
 \mathbf{v}(n+1) &= \mathbf{v}(n) + \mu \mathbf{x}(n) e(n) \\
 &= \mathbf{v}(n) - \mu \mathbf{R} \mathbf{v}(n) - \mu \mathbf{w}(n) \\
 &= (\mathbf{I} - \mu \mathbf{R}) \mathbf{v}(n) - \mu \mathbf{w}(n) \\
 &= (\mathbf{I} - \mu \mathbf{Q} \mathbf{D} \mathbf{Q}^H) \mathbf{v}(n) - \mu \mathbf{w}(n) \\
 &= \mathbf{Q} (\mathbf{I} - \mu \mathbf{D}) \mathbf{Q}^H \mathbf{v}(n) - \mu \mathbf{w}(n),
 \end{aligned} \tag{46}$$

where the covariance matrix can be decomposed as  $\mathbf{R} = \mathbf{Q} \mathbf{D} \mathbf{Q}^H$ . Here,  $\mathbf{Q}$  is an  $N \times N$  unitary matrix while  $\mathbf{D} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}$  is an  $N \times N$  eigenvalue diagonal matrix. We denote  $\tilde{\mathbf{v}}(n) = \mathbf{Q}^H \mathbf{v}(n)$  and  $\tilde{\mathbf{w}}(n) = \mathbf{Q}^H \mathbf{w}(n)$  as the rotated vectors, and Equation 46 can be rewritten as

$$\tilde{\mathbf{v}}(n+1) = (\mathbf{I} - \mu \mathbf{D}) \tilde{\mathbf{v}}(n) - \mu \tilde{\mathbf{w}}(n). \tag{47}$$

According to Equation 47, the MSE lower bound of LMS can be derived as

$$\begin{aligned}
 B &= \lim_{n \rightarrow \infty} E \{ \tilde{\mathbf{v}} \|\tilde{\mathbf{v}}(n+1)\|_2^2 \} \\
 &= \lim_{n \rightarrow \infty} E \{ [(\mathbf{I} - \mu \mathbf{D}) \tilde{\mathbf{v}}(n) - \mu \tilde{\mathbf{w}}^T(n)]^T [(\mathbf{I} - \mu \mathbf{D}) \tilde{\mathbf{v}}(n) - \mu \tilde{\mathbf{w}}(n)] \} \\
 &= \lim_{n \rightarrow \infty} \underbrace{E \{ \|\tilde{\mathbf{v}}(n)\|_2^2 \}}_{a(n)} - 2\mu \underbrace{E \{ \tilde{\mathbf{w}}^T(n) \tilde{\mathbf{v}}(n) \}}_{b(n)} \\
 &\quad + \underbrace{\mu^2 E \{ \tilde{\mathbf{w}}^T(n) \tilde{\mathbf{w}}(n) \}}_{c(n)}.
 \end{aligned} \tag{48}$$

Since signal and noise are independent, hence,  $E \{ \tilde{\mathbf{w}}^T(n) \tilde{\mathbf{v}}(n) \} = \dots = E \{ \tilde{\mathbf{w}}^T(n) \tilde{\mathbf{v}}(n) \} = 0$ , and Equation 48 can be simplified as

$$B = \lim_{n \rightarrow \infty} \underbrace{E \{ \|\tilde{\mathbf{w}}(n)\|_2^2 \}}_{a(n)} + \underbrace{\mu^2 E \{ \tilde{\mathbf{w}}^T(n) \tilde{\mathbf{w}}(n) \}}_{c(n)} \tag{49}$$

For a better understanding, the first term  $a(n)$  in Equation 49 can be expanded as

According to Equation 50, Equation 49 can be further rewritten as

$$\left\{ \begin{aligned}
 \underbrace{E \{ \|\tilde{\mathbf{v}}(n)\|_2^2 \}}_{a(n)} &= \underbrace{E \{ \|\tilde{\mathbf{v}}(n-1)\|_2^2 \}}_{a(n-1)} + \underbrace{\mu^2 E \{ \tilde{\mathbf{w}}^T(n) \tilde{\mathbf{w}}(n) \}}_{c(n-1)} \\
 \underbrace{E \{ \|\tilde{\mathbf{v}}(n-1)\|_2^2 \}}_{a(n-1)} &= \underbrace{E \{ \|\tilde{\mathbf{v}}(n-2)\|_2^2 \}}_{a(n-2)} + \underbrace{\mu^2 E \{ \tilde{\mathbf{w}}^T(n-1) \tilde{\mathbf{w}}(n-1) \}}_{c(n-2)} \\
 &\vdots \\
 \underbrace{E \{ \|\tilde{\mathbf{v}}(1)\|_2^2 \}}_{a(1)} &= \underbrace{E \{ \|\tilde{\mathbf{v}}(0)\|_2^2 \}}_{a(0)} + \underbrace{\mu^2 E \{ \tilde{\mathbf{w}}^T(1) \tilde{\mathbf{w}}(1) \}}_{c(0)}
 \end{aligned} \right. \tag{50}$$

$$\begin{aligned}
 B &= \lim_{n \rightarrow \infty} \underbrace{E \{ \|\tilde{\mathbf{v}}(0)\|_2^2 \}}_{a(0)} + \underbrace{\mu^2 \sum_{i=0}^{n-1} E \{ \tilde{\mathbf{w}}^T(i) \tilde{\mathbf{w}}(i) \}}_{c(0)+c(1)+\dots+c(n-1)} \\
 &\geq \lim_{n \rightarrow \infty} \mu^2 \sum_{i=0}^{n-1} E \{ \tilde{\mathbf{w}}^T(i) \tilde{\mathbf{w}}(i) \},
 \end{aligned} \tag{51}$$

where the first term  $\lim_{n \rightarrow \infty} E \{ \|\tilde{\mathbf{v}}(n)\|_2^2 \} \rightarrow 0$  when  $|1 - \mu \lambda_i| < 1$ . Consider the MSE lower bound of the  $i$ th channel taps  $\{b_i; i = 0, 1, \dots, N-1\}$ . We obtain

$$b_i = \lim_{n \rightarrow \infty} \mu^2 \sum_{i=0}^{N-1} (1 - \mu \lambda_i)^{2i} E \{ |\tilde{w}_i(n)|^2 \} = \frac{\mu P_0}{2 - \mu \lambda_i}, \tag{52}$$

where  $E \{ |\tilde{w}_i(n)|^2 \} = \lambda_i P_0$  and  $P_0$  denotes the gradient noise power. For any overall channel, since the LMS adaptive channel estimation method does not use the channel sparse structure information, the MSE lower bound should be cumulated from all of the channel taps. Hence, the lower bound  $B_{LMS}$  of LMS is given by

$$B = \sum_{i=0}^{N-1} b_i = \sum_{i=0}^{N-1} \frac{\mu P_0}{2 - \mu \lambda_i} \geq \sum_{i=0}^{N-1} \frac{\mu P_0}{2 - \mu \lambda_{\min}} = \frac{\mu N P_0}{2 - \mu \lambda_{\min}} \sim \mathcal{O}(N), \tag{53}$$

where  $N$  is the channel length of  $\mathbf{h}$ ,  $\{\lambda_i; i = 0, 1, \dots, N-1\}$  are eigenvalues of the covariance matrix  $\mathbf{R}$  and  $\lambda_{\min}$  is its minimal eigenvalue.  $\square$

**Theorem 4** For an  $N$ -length sparse channel vector  $\mathbf{h}$  which consists of  $K$  non-zero taps, if  $\mu$  satisfies  $0 < \mu < 2/\lambda_{\max}$ , then the MSE lower bound of the sparse LMS adaptive channel estimator is  $B_S = \mu P_0 K / (2 - \mu \lambda_{\min}) \sim \mathcal{O}(K)$ .

*Proof* From Equation 53, we can easily find that the MSE lower bound of the adaptive sparse channel estimator has a direct relationship with the number of non-zero channel coefficients, i.e.,  $K$ . Let  $\Omega$  denote the set of non-zero taps' position, that is,  $h_i \neq 0$ , for  $i \in \Omega$  and  $h_i = 0$  for others. We can then obtain the lower bound of the sparse LMS as



**Table 1 Simulation parameters of (N)LMS-based algorithms**

	$\mu$	$\Upsilon_{ZA(N)}$	$\Upsilon_{RZA(N)}$	$\Upsilon_{LP(N)}$	$\Upsilon_{Lo(N)}$
LMS-based	0.05	$0.02\sigma_n^2$	$0.05\sigma_n^2$	$0.005\sigma_n^2$	$0.02\sigma_n^2$
NLMS-based	0.1	$0.002\sigma_n^2$	$0.005\sigma_n^2$	$0.0005\sigma_n^2$	$0.002\sigma_n^2$

$$\begin{aligned}
 B_S &= \sum_{i=0, i \in \Omega}^{N-1} b_i = \sum_{i \in \Omega} \frac{\mu P_0}{2 - \mu \lambda_i} \geq \sum_{i \in \Omega} \frac{\mu P_0}{2 - \mu \lambda_{\min}} \\
 &= \frac{\mu K P_0}{2 - \mu \lambda_{\min}} \sim \mathcal{O}(K).
 \end{aligned} \tag{54}$$

□

#### 4 Computer simulations

In this section, we will compare the performance of the proposed channel estimators using 1,000 independent Monte Carlo runs for averaging. The length of sparse multipath channel  $\mathbf{h}$  is set as  $N = 16$ , and its number of dominant taps is set as  $K = 1, 2, 4$ , and  $8$ . For simplicity, the number of dominant taps is written in the form of a set as  $K \in \{1, 2, 4, 8\}$ . The values of the dominant channel taps follow the Gaussian distribution, and the positions of dominant taps are randomly selected within the length of  $\mathbf{h}$  and is subjected to  $\|\mathbf{h}\|_2^2 = 1$ . The received signal-to-noise ratio (SNR) is defined as  $10 \log(E_0/\sigma_n^2)$ , where  $E_0 = 1$  is the received signal power and the noise power is

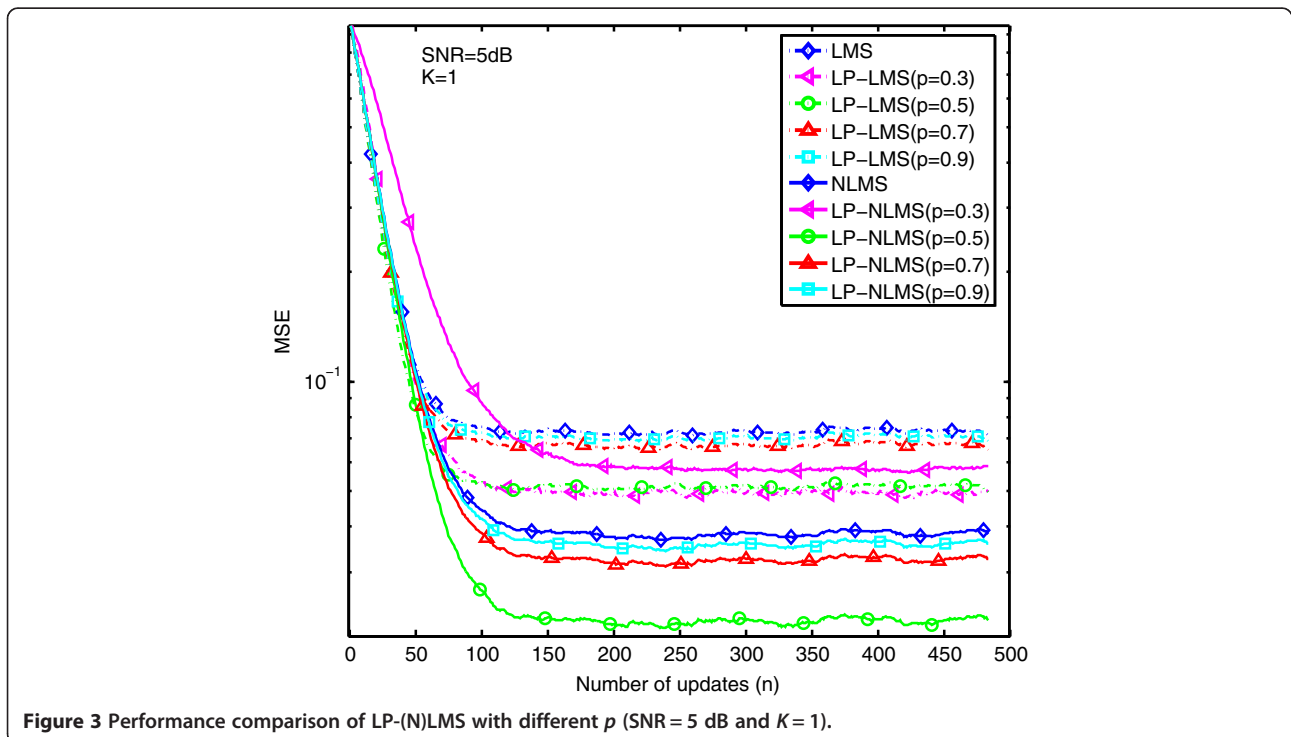
given by  $\sigma_n^2 = 10^{-\text{SNR}/10}$ . Here, we compare their performance with three SNR regimes:  $\{5, 10, 20 \text{ dB}\}$ . All of the step sizes and regularization parameters are listed in Table 1. It is worth noting that the (N)LMS-based algorithms can exploit more accurate sparse channel information at higher SNR environment. Hence, all of the parameters are set at a direct ratio in relation with noise power. For example, in the case of SNR = 10 dB, the parameters of LMS-based algorithm are matched with the parameters which are given in [20]. Hence, the propose regulation parameter method can adaptively exploit the channel sparsity under different SNR regimes.

The estimation performance between the actual and estimated channels is evaluated by MSE standard which is defined as

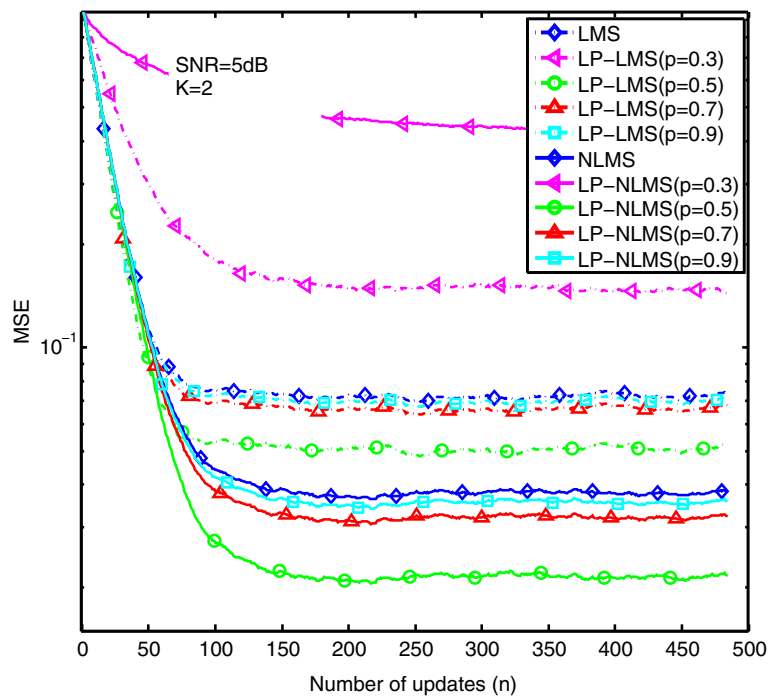
$$\text{MSE}\{\tilde{\mathbf{h}}(n)\} = E\left\{\|\mathbf{h} - \tilde{\mathbf{h}}(n)\|_2^2\right\}, \tag{55}$$

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

In the first experiment, we evaluate the estimation performance of LP-(N)LMS as a function of  $p \in \{0.3, 0.5, 0.7, 0.9\}$  which are shown in Figures 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, and 14 in three SNR regimes, i.e.,  $\{5, 10, 20 \text{ dB}\}$ . The parameter is set as  $\varepsilon_{LP} = \varepsilon_{LPN} = 0.05$  which is suggested in [20]. In the case of low SNR regime, e.g., SNR =



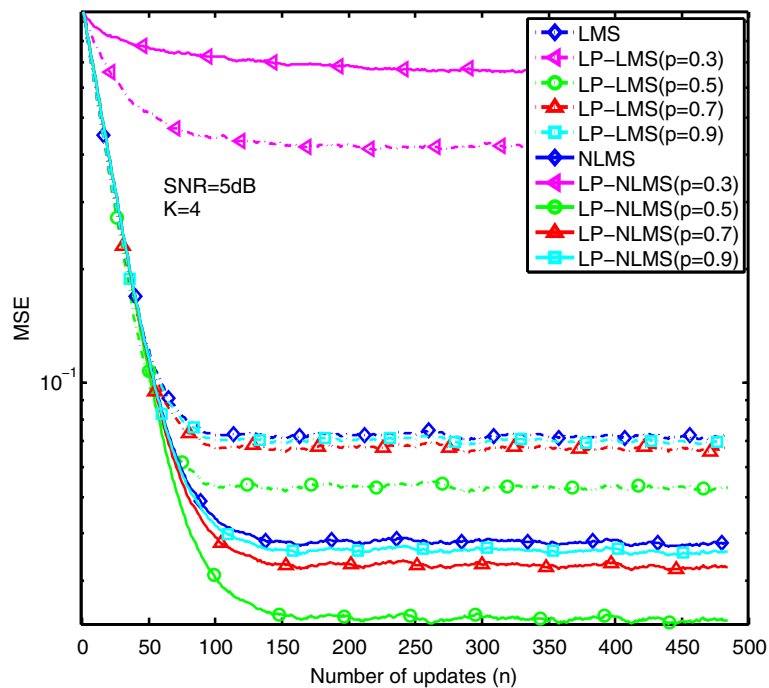
**Figure 3 Performance comparison of LP-(N)LMS with different  $p$  (SNR = 5 dB and  $K = 1$ ).**



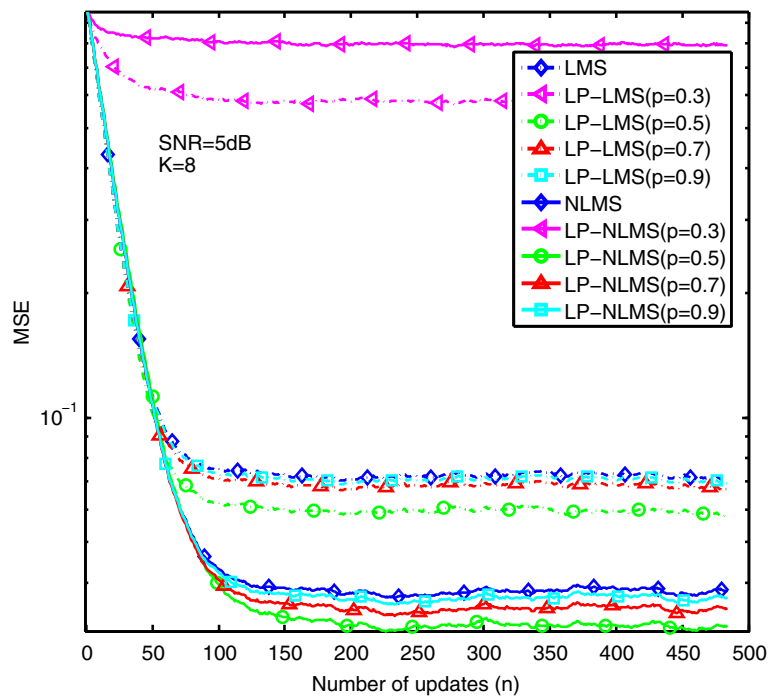
**Figure 4** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 5 dB and  $K = 2$ ).

5 dB, MSE performance curves are depicted in Figures 3, 4, 5, and 6, with different  $K \in \{1, 2, 4, 8\}$ . One can find that LP-(N)LMS algorithm is not stable if we set  $p = 0.3$  because their estimation performances are even worse than the other cases, i.e.,  $p \in \{0.5, 0.7, 0.9\}$ . In the case of

intermediate SNR regime, e.g., SNR = 10 dB, MSE performance curves are depicted in Figures 7, 8, 9, and 10, with different  $K \in \{1, 2, 4, 8\}$ . As shown in Figure 7, to estimate a very sparse channel ( $K = 1$ ), the LP-(N)LMS algorithm with  $p = 0.3$  can obtain a better estimation



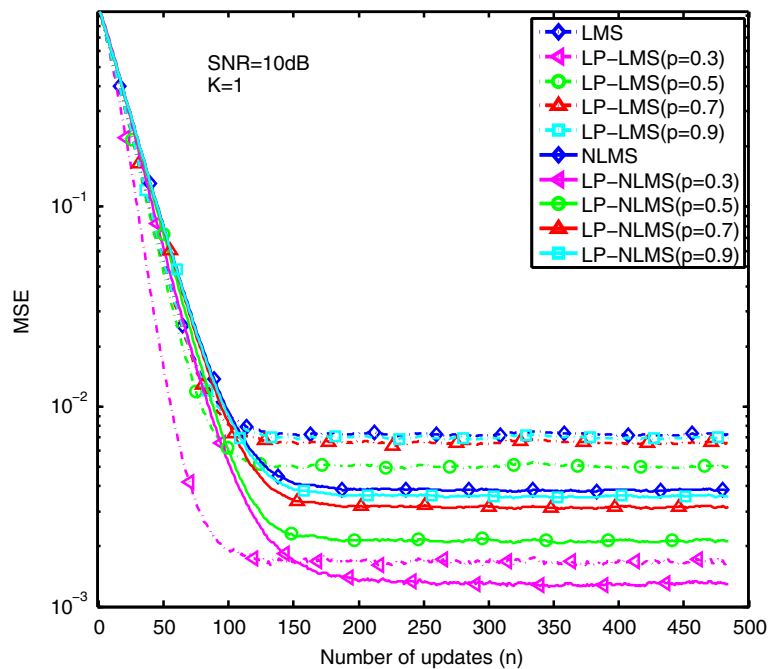
**Figure 5** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 5 dB and  $K = 4$ ).



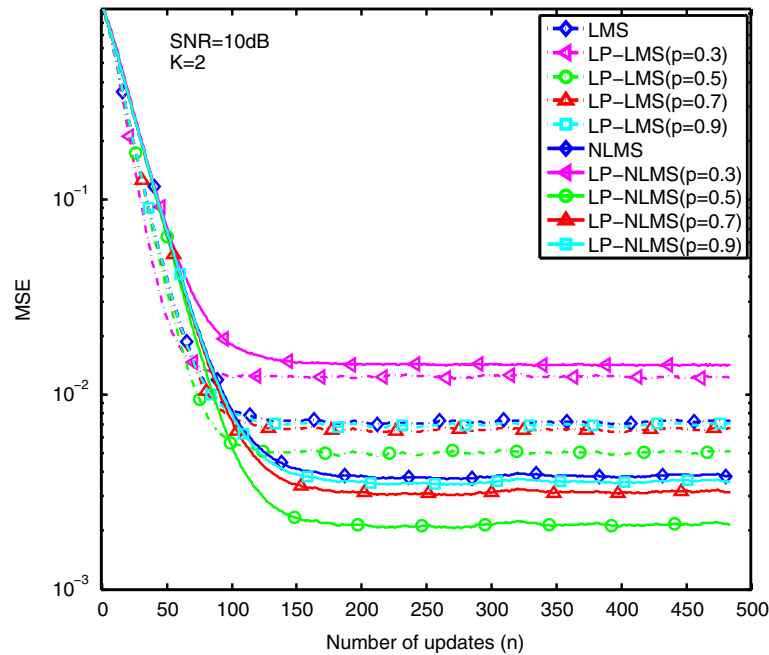
**Figure 6** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 5 dB and  $K = 8$ ).

performance than the others, i.e.,  $p \in \{0.5, 0.7, 0.9\}$ . However, if  $K > 1$ , the LP-(N)LMS algorithm is no longer stable as shown in Figures 8, 9, and 10, whose estimation performance is even worse than that in (N)LMS. In the case of high SNR regime, e.g., SNR = 20 dB,

likewise, the LP-(N)LMS algorithm achieves a better estimation performance than the others only when  $K \leq 2$  as shown in Figures 11 and 12. Hence, one can find that the stability of the LP-(N)LMS algorithm with  $p = 0.3$  depends highly on both SNR and  $K$ . If the algorithm adopts



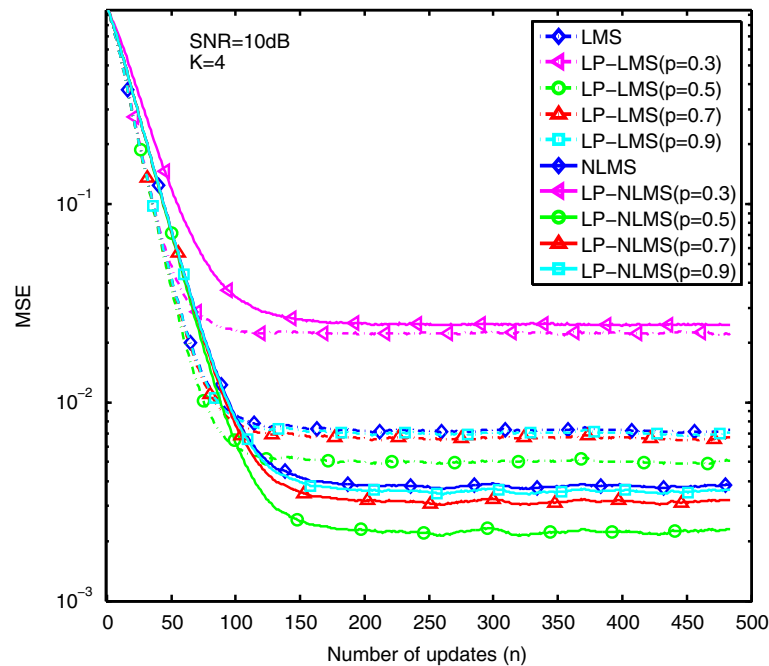
**Figure 7** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 10 dB and  $K = 1$ ).



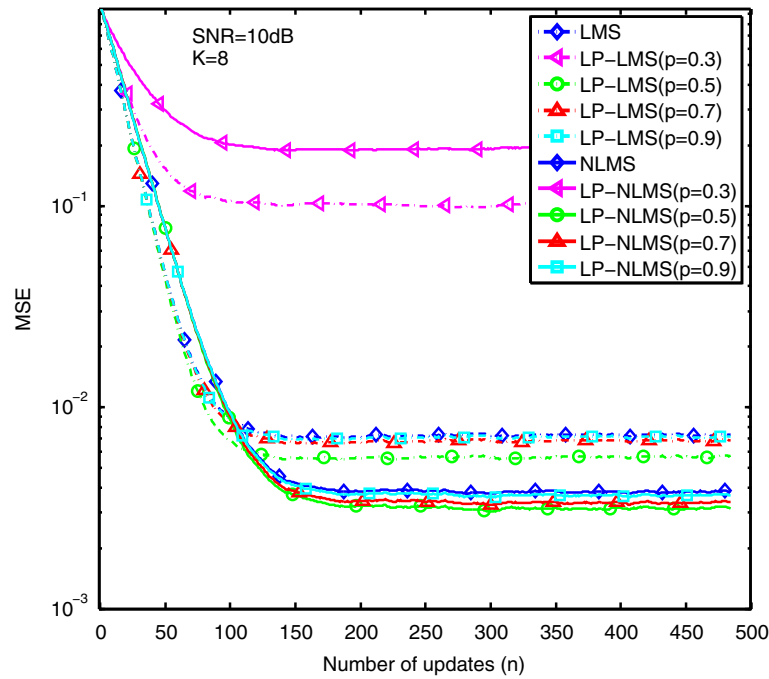
**Figure 8** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 10 dB and  $K = 2$ ).

$p \in \{0.5, 0.7, 0.9\}$  as shown in Figures 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, and 14, fortunately, there is no obvious relationship between stability and either SNR or  $K$ . In order to trade off stability and estimation performance of the LP-(N)LMS algorithm, it is better to set the value of the sparse penalty as  $p = 0.5$ . For one thing, the LP-(N)LMS

algorithm with  $p = 0.5$  can always achieve a better estimation performance than the cases with  $p \in \{0.7, 0.9\}$ . In another case, even though the LP-(N)LMS algorithm with  $p = 0.3$  can obtain a better estimation performance in certain circumstances, its stability depends highly on SNR and  $K$ . Hence, in the following



**Figure 9** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 10 dB and  $K = 4$ ).

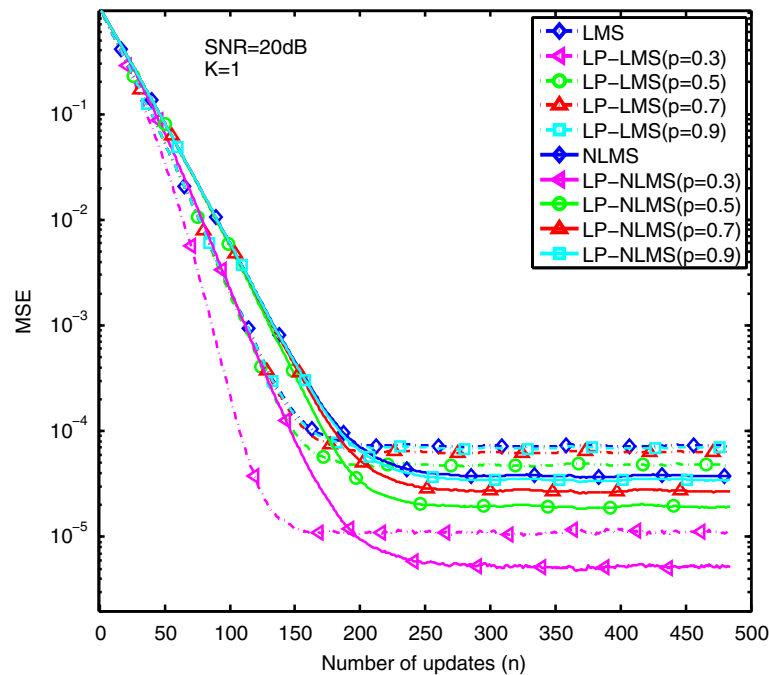


**Figure 10** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 10 dB and  $K = 8$ ).

simulation results,  $p = 0.5$  is considered for LP-(N) LMS-based adaptive sparse channel estimation.

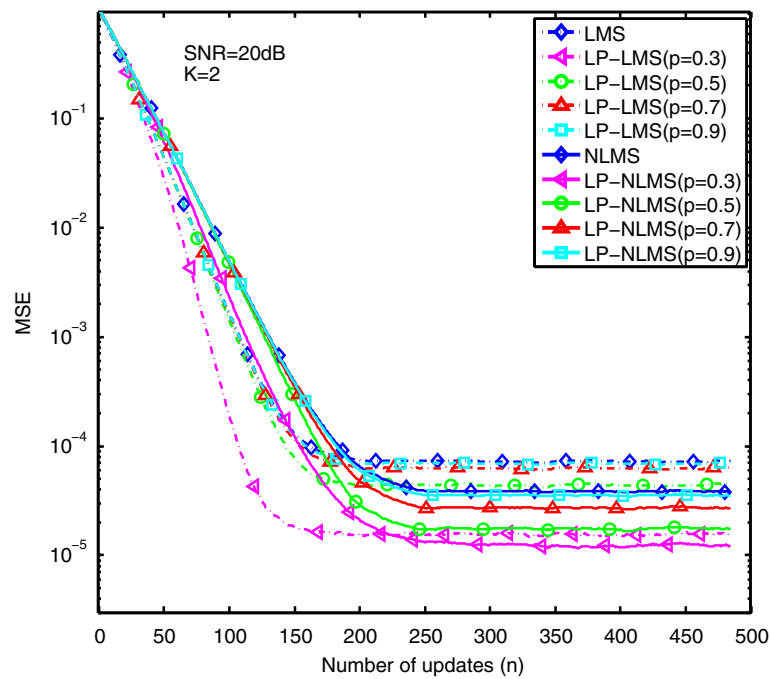
In the second experiment, we compare all the (N) LMS-based sparse adaptive channel estimation methods with different SNR regimes: {5, 20 dB} as

well as  $K \in \{1, 8\}$ , as shown in Figures 15, 16, 17, and 18, respectively. In the case of low SNR regime (e.g., SNR = 5 dB), the MSE curves show that NLMS-based methods achieve a better estimation performance than LMS-based ones as shown in Figure 15. Let us take ZA-NLMS and



**Figure 11** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 20 dB and  $K = 1$ ).

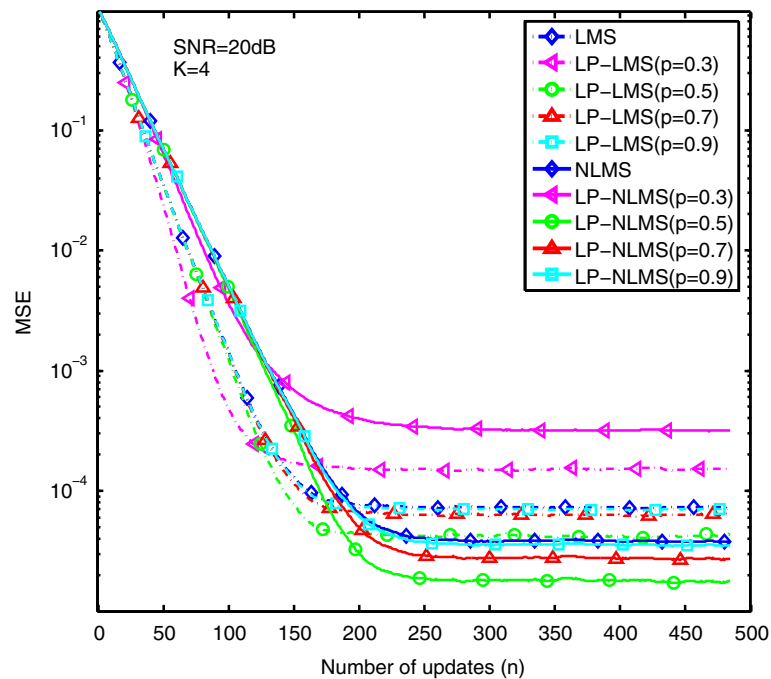




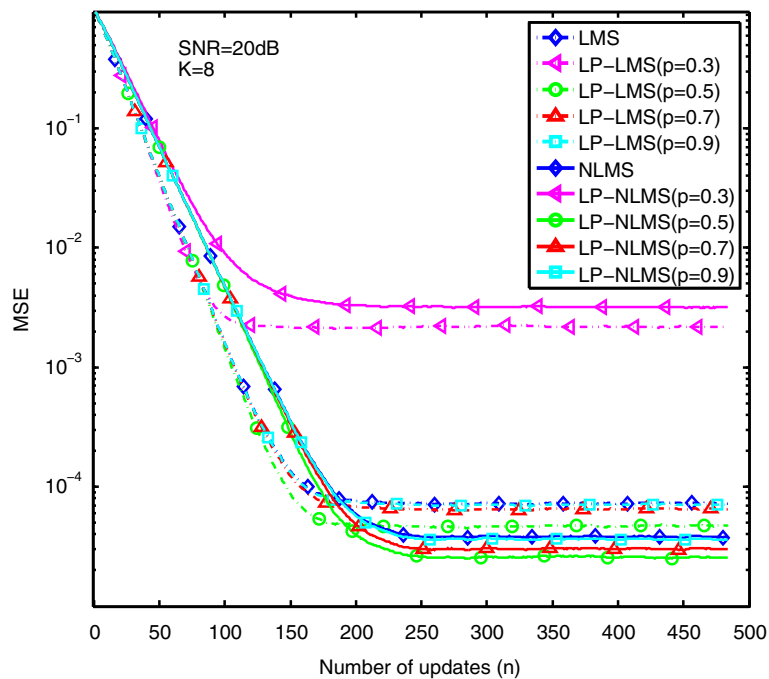
**Figure 12** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 20 dB and  $K = 2$ ).

ZA-LMS as examples. As shown in the figure, each performance curve of ZA-NLMS is much lower than that of ZA-LMS. That is to say, the proposed ZA-NLMS achieves a better estimation performance than the traditional ZA-LMS. Similarly, other proposed NLMS-type methods, i.e.,

RZA-NLMS and LP-NLMS, also achieve better estimation performances than their corresponding sparse LMS types. To further confirm the stability and effectiveness of our proposed methods, sparse (N) LMS-based estimation methods are also evaluated in



**Figure 13** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 20 dB and  $K = 4$ ).



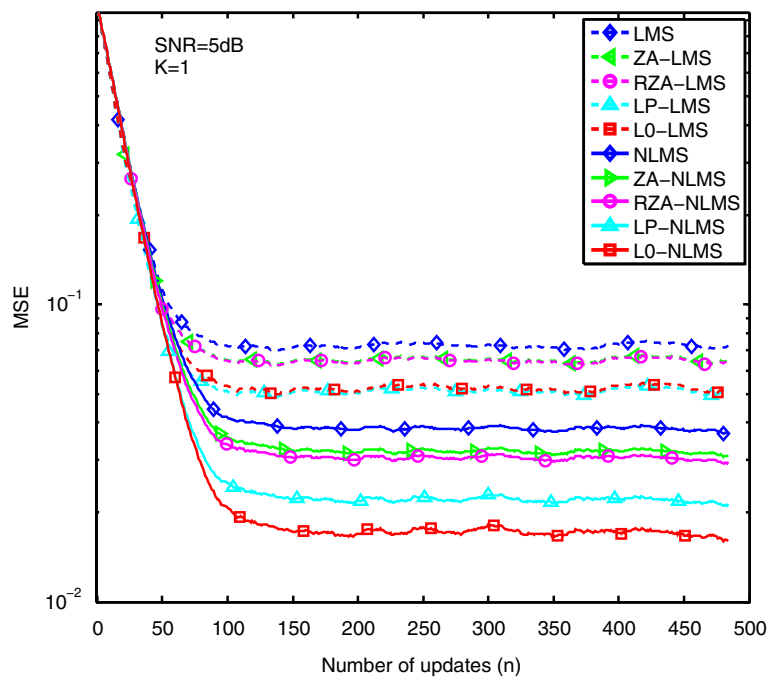
**Figure 14** Performance comparison of LP-(N)LMS with different  $p$  (SNR = 20 dB and  $K = 8$ ).

the case of 20 dB as well as with different  $K$ , respectively.

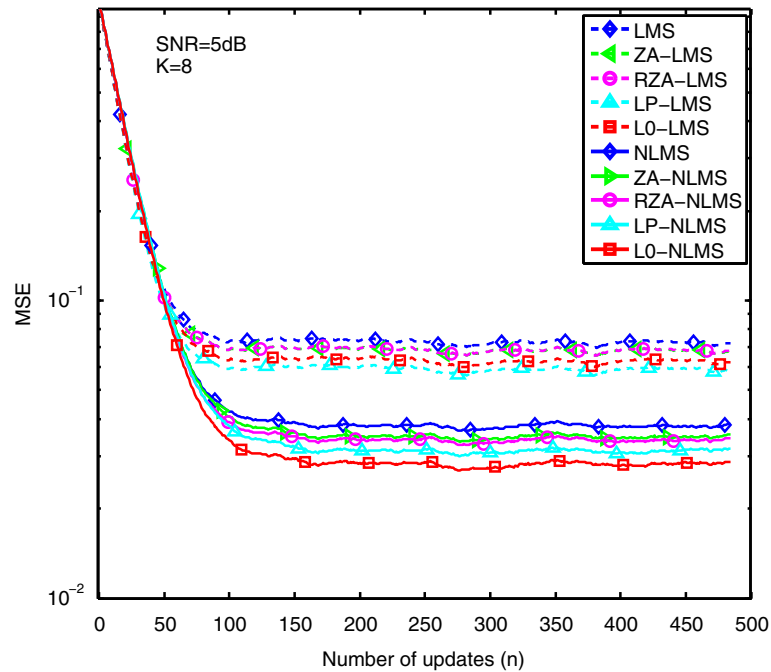
As the SNR increases, the obvious performance advantages of NLMS-based methods are vanishing. Hence, compared with LMS-based methods, we can conclude

that NLMS-based methods not only work more reliably for unknown signal scaling, but also work more stably for noise interference, especially in low SNR environment.

In addition, the simulation results also show that (N) LMS-based methods have an inverse relationship with



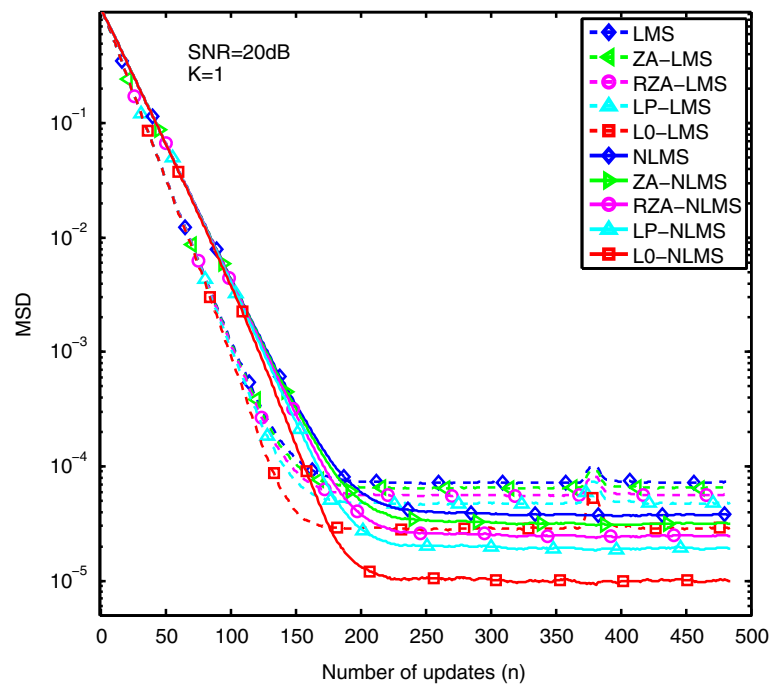
**Figure 15** MSE versus the number of iterations (SNR = 5 dB and  $K = 1$ ).



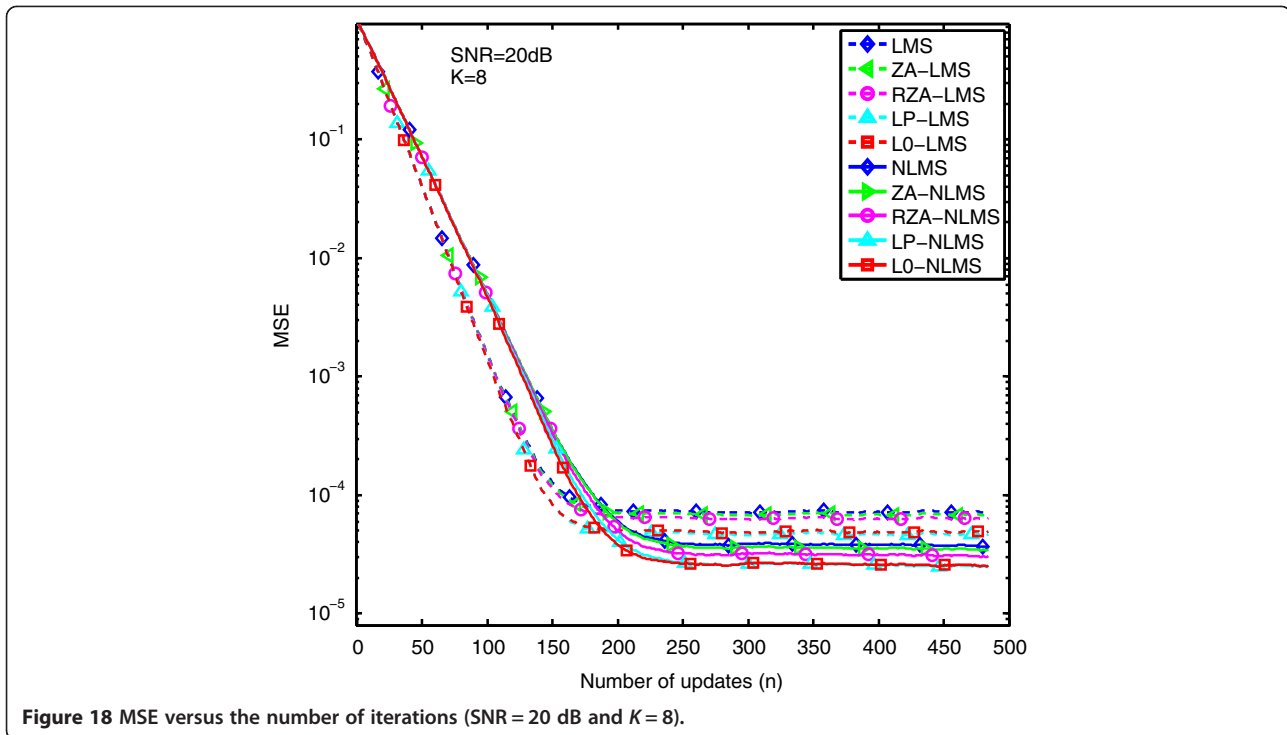
**Figure 16** MSE versus the number of iterations (SNR = 5 dB and  $K = 8$ ).

the number of non-zero channel taps. In other words, for a sparse channel, (N)LMS-based methods can achieve a better estimation performance and vice versa. Let us take  $K = 1$  and 8 as examples. In Figures 15 and 16, when the number of non-zero taps is  $K = 1$ , performance gaps are bigger

than in the case where  $K = 8$ , as shown in Figures 17 and 18. Hence, these simulation results also show that the estimation performance of adaptive sparse channel estimation is also affected by the number of non-zero channel taps. When the channel is no longer sparse, the performance of



**Figure 17** MSE versus the number of iterations (SNR = 20 dB and  $K = 1$ ).



**Figure 18** MSE versus the number of iterations (SNR = 20 dB and K = 8).

these proposed methods will reduce to the performance of (N)LMS-based methods.

## 5 Conclusions

In this paper, we have investigated various (N)LMS-based adaptive sparse channel estimation methods by enforcing different sparse penalties, e.g.,  $\ell_p$ -norm and  $\ell_0$ -norm. The research motivation originated from the fact that LMS-based channel estimation methods are sensitive to the scaling of random training signal and easily causing the estimation performance unstable. Unlike LMS-based methods, the proposed NLMS-based methods have avoided the uncertain signal scaling and normalized the power of input signal with different sparse penalties.

Initially, we proposed an improved adaptive sparse channel estimation method using  $\ell_0$ -norm sparse constraint LMS algorithm and compared it with ZA-LMS, RZA-LMS, and LP-LMS. The proposed method was based on the CS background that  $\ell_0$ -norm sparse penalty can exploit a more accurate channel sparsity.

In addition, to improve the robust performance and increase the convergence speed, we proposed NLMS-based adaptive sparse channel estimation methods using different sparse penalties, i.e., ZA-NLMS, RZA-NLMS, LP-NLMS, and L0-NLMS. For example, ZA-NLMS can achieve a better estimation than ZA-LMS. The proposed methods exhibit faster convergence and better performance which are confirmed by computer simulations under various SNR environments.

## Competing interests

The authors declare that they have no competing interests.

## Acknowledgements

The authors would like to thank Dr. Koichi Adachi of the Institute for Infocomm Research for his valuable comments and suggestions as well as for the improvement of the English expression of this paper. The authors would like to extend their appreciation to the anonymous reviewers for their constructive comments. This work was supported by a grant-in-aid for the Japan Society for the Promotion of Science (JSPS) fellows (grant number 24-02366).

Received: 20 December 2012 Accepted: 28 July 2013

Published: 5 August 2013

## References

1. D Raychaudhuri, N Mandayam, Frontiers of wireless and mobile communications. *Proc. IEEE* **100**(4), 824–840 (2012)
2. F Adachi, D Grag, S Takaoka, K Takeda, Broadband CDMA techniques. *IEEE Wirel. Commun.* **12**(2), 8–18 (2005)
3. F Adachi, E Kudoh, New direction of broadband wireless technology. *Wirel. Commun. Mob. Com.* **7**(8), 969–983 (2007)
4. WF Schreiber, Advanced television systems for terrestrial broadcasting: some problems and some proposed solutions. *Proc. IEEE* **83**(6), 958–981 (1995)
5. AF Molisch, Ultra wideband propagation channels: theory, measurement, and modelling. *IEEE Trans. Veh. Technol.* **54**(5), 1528–1545 (2005)
6. Z Yan, M Herdin, AM Sayeed, E Bonek, *Experimental study of MIMO channel statistics and capacity via the virtual channel representation* (University of Wisconsin-Madison, Technique report, 2007)
7. N Czik, X Yin, H Ozcelik, M Herdin, E Bonek, BH Fleury, Cluster characteristics in a MIMO indoor propagation environment. *IEEE Trans. Wirel. Commun.* **6**(4), 1465–1475 (2007)
8. L Vuokko, VM Kolmonen, J Salo, P Vainikainen, Measurement of large-scale cluster power characteristics for geometric channel models. *IEEE Trans. Antennas Propagat.* **55**(11), 3361–3365 (2007)
9. B Widrow, SD Stearns, *Adaptive Signal Processing* (Prentice Hall, New Jersey, 1985)
10. E Candes, J Romberg, T Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory* **52**(2), 489–509 (2006)
11. DL Donoho, Compressed sensing. *IEEE Trans. Inform. Theory* **52**(4), 1289–1306 (2006)

12. WU Bajwa, J Haupt, AM Sayeed, R Nowak, Compressed channel sensing: a new approach to estimating sparse multipath channels. *Proc. IEEE* **98**(6), 1058–1076 (2010)
13. G Taubock, F Hlawatsch, D Eiwien, H Rauhut, Compressive estimation of doubly selective channels in multicarrier systems: leakage effects and sparsity-enhancing processing. *IEEE J. Sel. Top. Sign. Process.* **4**(2), 255–271 (2010)
14. G Gui, Q Wan, W Peng, F Adachi, *Sparse multipath channel estimation using compressive sampling matching pursuit algorithm. The 7th IEEE Vehicular Technology Society Asia Pacific Wireless Communications Symposium (APWCS), Kaohsiung, 20–21 May 2010* pp. 1–5
15. R Tibshirani, Regression shrinkage and selection via the Lasso. *J. Roy. Stat. Soc. (B)* **58**(1), 267–288 (1996)
16. EJ Candes, The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique* **346**(9–10), 589–592 (2008)
17. MR Garey, DS Johnson, *Computers and Intractability: a Guide to the Theory of NP-Completeness* (W.H. Freeman & Co, New York, NY, 1990)
18. Y Chen, Y Gu, AO Hero, *Sparse LMS for system identification. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Taipei, 19–24 April 2009* pp. 3125–3128
19. EJ Candes, MB Wakin, SP Boyd, Enhancing sparsity by reweighted  $\ell_1$ -minimization. *J. Fourier Anal. Appl.* **14**(5–6), 877–905 (2008)
20. O Taheri, SA Vorobyov, *Sparse channel estimation with  $\ell_p$ -norm and reweighted  $\ell_1$ -norm penalized least mean squares. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, 22–27 May 2011* pp. 2864–2867
21. S Haykin, *Adaptive Filter Theory*, 3rd edn. (Prentice-Hall, Englewood Cliffs, 1996)
22. YT Gu, J Jin, S Mei,  $\ell_0$ -Norm constraint LMS algorithm for sparse system identification. *IEEE Signal Process. Lett.* **16**(9), 774–777 (2009)
23. GL Su, J Jin, YT Gu, J Wang, Performance analysis of  $\ell_0$  norm constraint least mean square algorithm. *IEEE Trans. Signal Process.* **60**(5), 2223–2235 (2012)
24. G Gui, W Peng, F Adachi, *Improved adaptive sparse channel estimation based on the least mean square algorithm. IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, 7–10 April 2013* pp. 3130–3134
25. G Gui, A Mehbodniya, F Adachi, *Least mean square/fourth algorithm for adaptive sparse channel estimation. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), London, 8–11 September 2013.* in press
26. G Gui, A Mehbodniya, F Adachi, *Adaptive sparse channel estimation using re-weighted zero-attracting normalized least mean fourth. 2nd IEEE/CIC International Conference on Communications in China (ICCC), Xian, 12–14 August 2013.* in press
27. Z Huang, G Gui, A Huang, D Xiang, F Adachi, *Regularization selection method for LMS-type sparse multipath channel estimation, in 19th Asia-Pacific Conference on Communications (APCC), Bali, 29–31 August 2013.* in press

doi:10.1186/1687-1499-2013-204

**Cite this article as:** Gui and Adachi: Improved least mean square algorithm with application to adaptive sparse channel estimation. *EURASIP Journal on Wireless Communications and Networking* 2013 **2013**:204.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)

---