

## Article

# Sign Function Based Sparse Adaptive Filtering Algorithms for Robust Channel Estimation under Non-Gaussian Noise Environments

Tingping Zhang <sup>1,2,\*</sup> and Guan Gui <sup>3</sup><sup>1</sup> School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China<sup>2</sup> College of Computer Science, Chongqing University, Chongqing 400044, China<sup>3</sup> Institute of Signal Transmission and Processing, College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; guiguan@njupt.edu.cn

\* Correspondence: ztp@cqjtu.edu.cn; Tel./Fax: +86-23-6265-2751

Academic Editor: Paul M. Goggans

Received: 24 June 2016; Accepted: 9 August 2016; Published: 12 August 2016

**Abstract:** Robust channel estimation is required for coherent demodulation in multipath fading wireless communication systems which are often deteriorated by non-Gaussian noises. Our research is motivated by the fact that classical sparse least mean square error (LMS) algorithms are very sensitive to impulsive noise while standard SLMS algorithm does not take into account the inherent sparsity information of wireless channels. This paper proposes a sign function based sparse adaptive filtering algorithm for developing robust channel estimation techniques. Specifically, sign function based least mean square error (SLMS) algorithms to remove the non-Gaussian noise that is described by a symmetric  $\alpha$ -stable noise model. By exploiting channel sparsity, sparse SLMS algorithms are proposed by introducing several effective sparse-promoting functions into the standard SLMS algorithm. The convergence analysis of the proposed sparse SLMS algorithms indicates that they outperform the standard SLMS algorithm for robust sparse channel estimation, which can be also verified by simulation results.

**Keywords:** robust sparse channel estimation; sign function based least mean square error (SLMS); sparsity-promoting function; non-Gaussian noise; convergence analysis

## 1. Introduction

Broadband signal transmission is considered an indispensable technique in next-generation dependable wireless communication systems [1–3]. It is well known that both multipath fading and additive noises are major determinants that impair the system performance. In such circumstances, either coherence detection or demodulation needs to estimate channel state information (CSI) [1]. In the framework of a Gaussian noise model, some effective channel estimation techniques have been studied [4–10]. In the assumptions of the non-Gaussian impulsive noise model, however, existing estimation techniques do not perform robustly due to heavy tailed impulsive interference. Generally speaking, impulsive noise is used to generate natural or man-made electromagnetic waves that are different from the conventional Gaussian noise model [11]. For example, a second-order statistics-based least mean square error (LMS) algorithm [4] cannot be directly applied in broadband channel estimation [12]. To solve this problem, selecting a suitable noise model is necessary to devise a stable channel estimation that can combat the harmful impulsive noises.

The aforementioned non-Gaussian noise can be modeled by the symmetric alpha-stable (S $\alpha$ S) distribution [13]. Based on the S $\alpha$ S noise model, several adaptive filtering based robust channel estimation techniques have been developed [14–16]. These techniques are based on the channel model

assumption of dense finite impulse response (FIR), which may not be suitable in broadband channel estimation because the channel vector is supported only by a few dominant coefficients [17,18].

Considering the sparse structure in a wireless channel, this paper proposes a kind of sparse SLMS algorithm with different sparse norm constraint functions. Specifically, we adopt five sparse constraint functions as follows: zero-attracting (ZA) [7] and reweighted ZA (RZA) [7], reweighted  $\ell_1$ -norm (RL1) [19],  $\ell_p$ -norm (LP), and  $\ell_0$ -norm (L0) [20], to take advantage of sparsity and to mitigate non-Gaussian noise interference. It is necessary to state that the short versions of the proposed algorithms were initially presented in a conference but we did not give a performance analysis [21]. In this paper, we first propose five sparse SLMS algorithms for channel estimation. To verify the stability of the proposed SLMS algorithms, convergence analysis is derived with respect to mean and excess mean square error (MSE) performance. Finally, numerical simulations are provided to verify the effectiveness of the proposed algorithms.

The rest of the paper is organized as follows. Section 2 introduces an alpha-stable impulsive noise based sparse system model and traditional channel estimation technique. Based on the sparse channel model, we propose five sparse SLMS algorithms in Section 3. To verify the proposed sparse SLMS algorithms, convergence analysis is derived in Section 4. Later, computer simulations are provided to validate the effectiveness of the proposed algorithms in Section 5. Finally, Section 6 concludes the paper and proposes future work.

## 2. Traditional Channel Estimation Technique

An input–output wireless system under the S $\alpha$ S noise environment is considered. The wireless channel vector is described by  $N$ -length FIR sparse vector  $\mathbf{w} = [w_0, w_1, \dots, w_{N-1}]^T$  at discrete time-index  $n$ . The received signal is obtained as

$$d(n) = \mathbf{w}^T \mathbf{x}(n) + z(n) \quad (1)$$

where  $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$  is the input signal vector of the  $N$  most recent input samples with distribution of  $\mathcal{CN}(0, \sigma_x^2)$  and  $z(n) \sim \phi(\alpha, \beta, \gamma, \delta)$  denotes a S $\alpha$ S noise variable. To understand the characteristic function of S $\alpha$ S noise, here we define it as

$$p(t) = \exp \{j\delta t - \gamma^\alpha |t|^\alpha [1 - j\beta \text{sgn}(t)\phi(t, \alpha)]\} \quad (2)$$

where

$$0 < \alpha \leq 2, -1 \leq \beta \leq 1, \gamma > 0, -\infty < \delta < \infty \quad (3)$$

$$\text{sgn}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases} \quad (4)$$

$$\phi(t, \alpha) = \begin{cases} \tan(\alpha\pi/2), & \alpha \neq 1 \\ -\log(|t|), & \alpha = 1 \end{cases} \quad (5)$$

In Equation (2),  $\alpha \in (0, 2]$  controls the tail heaviness of S $\alpha$ S noise. Since when  $\alpha < 1$  is rare to happen S $\alpha$ S noise in practical systems,  $\alpha \in (1, 2]$  is considered throughout this paper [11].  $\gamma > 0$  denotes the dispersive parameter, which can perform a similar role to Gaussian distribution;  $\beta \in [-1, 1]$  stands for the symmetrical parameter. To have a better understanding of the alpha-stable noise, its probability density function (PDF) curves are depicted in Figures 1 and 2 as examples. Figure 1a shows that the PDF curve of symmetric alpha-stable noise changes with the parameter  $\alpha$ , i.e., a smaller  $\alpha$  produces a larger PDF of the alpha-stable noise model and vice versa. In other words,  $\alpha$  controls the strength of the impulsive noise. Similarly, Figure 1b shows that the PDF curve of skewed  $\alpha$ -stable noise model also changes simultaneously with  $\alpha$  and  $\beta$ . Since the skewed noise model may not exist in

practical wireless communication systems, the symmetrical  $\alpha$ -stable noise model is used in this paper; the characteristic function of  $\alpha$ -stable process reduces as

$$p(t) = \exp(-\gamma |t|^\alpha) \quad (6)$$

For convenience, symmetric  $\alpha$ -stable noise variance is defined by

$$\sigma_z^2 = \gamma^{2/\alpha} \quad (7)$$

and the generalized received signal noise ratio (GSNR) is defined by

$$E_s/N_0 \text{ (dB)} \triangleq 10 \times \log_{10} \left\{ P_0/\gamma^{2/\alpha} \right\} \quad (8)$$

where  $P_0$  denotes the received signal power while  $\sigma_z^2 = \gamma^{2/\alpha}$  plays the same role as the noise variance.

The objective of adaptive channel estimation is to perform adaptive estimate of  $w(n)$  with limited complexity and memory given sequential observation  $\{d(n), x(n)\}$  in the presence of additive noise  $z(n)$ . That is to say, the estimate observation signal  $y(n)$  is given as

$$y(n) = w^T(n)x(n) \quad (9)$$

where  $w(n)$  denotes channel estimator. By combining (1) and (4), the estimation error  $e(n)$  is

$$e(n) = d(n) - y(n) = z(n) - x^T(n)v(n) \quad (10)$$

where  $v(n) = w(n) - w$  is the updating error of  $w(n)$  at iteration  $n$ . The cost function of standard LMS was written as

$$G_{LMS}(w(n)) = (1/2) e^2(n) \quad (11)$$

Using Equation (11), the standard LMS algorithm was derived as

$$\begin{aligned} w(n+1) &= w(n) + \mu \frac{\partial G(w(n))}{\partial w(n)} \\ &= w(n) + \mu e(n)x(n) \end{aligned} \quad (12)$$

where  $\mu$  denotes a step-size that controls the gradient descend speed of the LMS. Letting  $R = E[x(n)x^T(n)]$  denotes the covariance matrix of input signal  $x(n)$  and  $\lambda_{\max}$  as its maximum eigenvalue. The well-known stable convergence condition of the SLMS is

$$0 < \mu_{LMS} < 1/\lambda_{\max} \quad (13)$$

In order to remove S&S noise, the traditional SLMS algorithm [14] was first proposed as

$$w(n+1) = w(n) + \mu \text{sgn}(e(n)) x(n) \quad (14)$$

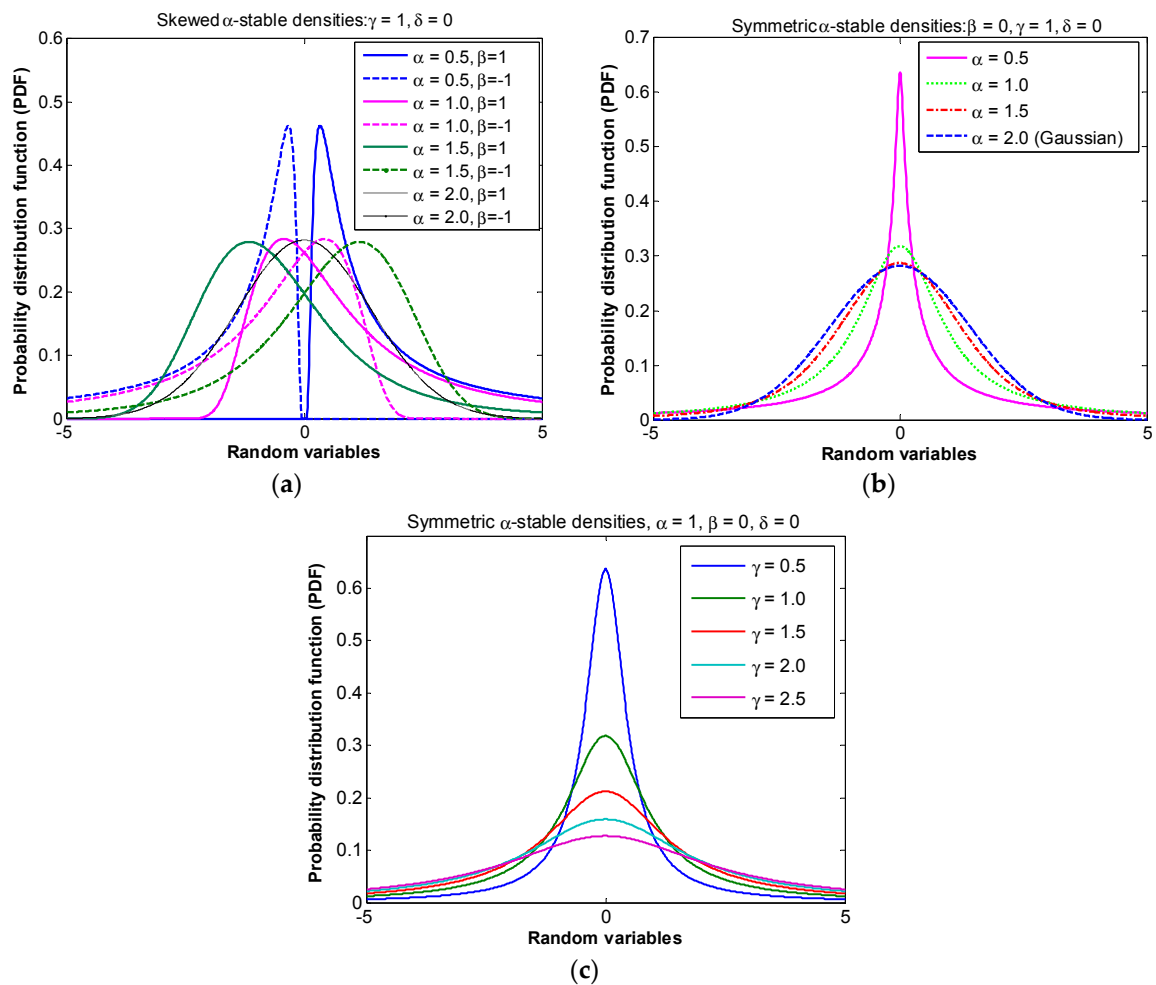
To ensure the stability of SLMS,  $\mu$  should be chosen as

$$0 < \mu < \sqrt{2\pi}\sigma_e(n)/\lambda_{\max} \quad (15)$$

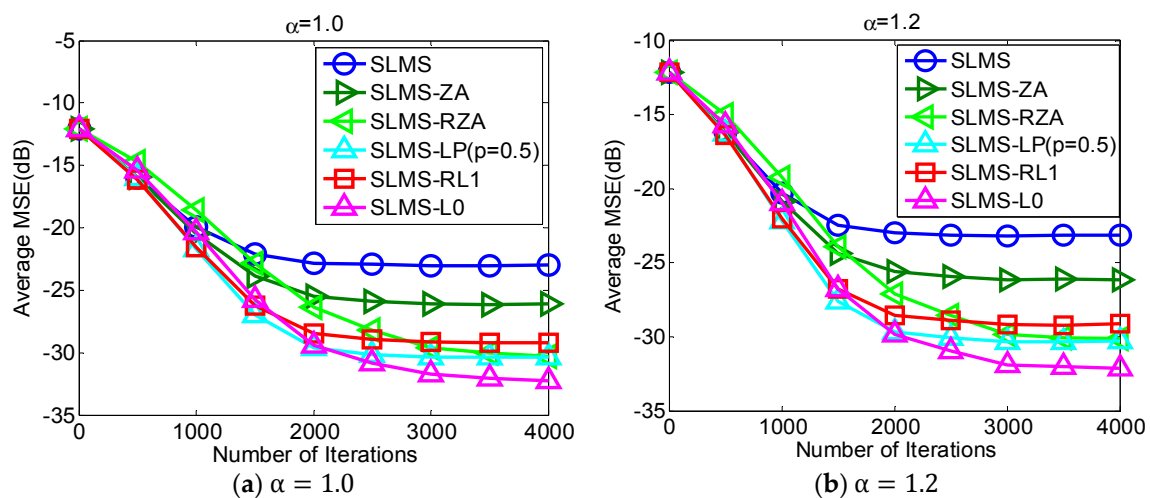
where  $\sigma_e(n)$  denotes the unconditional variance of estimation error  $e(n)$ . For later theoretical analysis, the  $\sigma_e^2$  is conditioned by

$$\begin{aligned} E\{e^2(n)|v(n)\} &\approx E\{e^2(n)\} = \sigma_e^2(n) \\ &= E\left\{[z(n) - v^T(n)x(n)]^T [z(n) - v^T(n)x(n)]\right\} \\ &= \gamma^{2/\alpha} + \text{Tr}\{RC(n)\} \end{aligned} \quad (16)$$

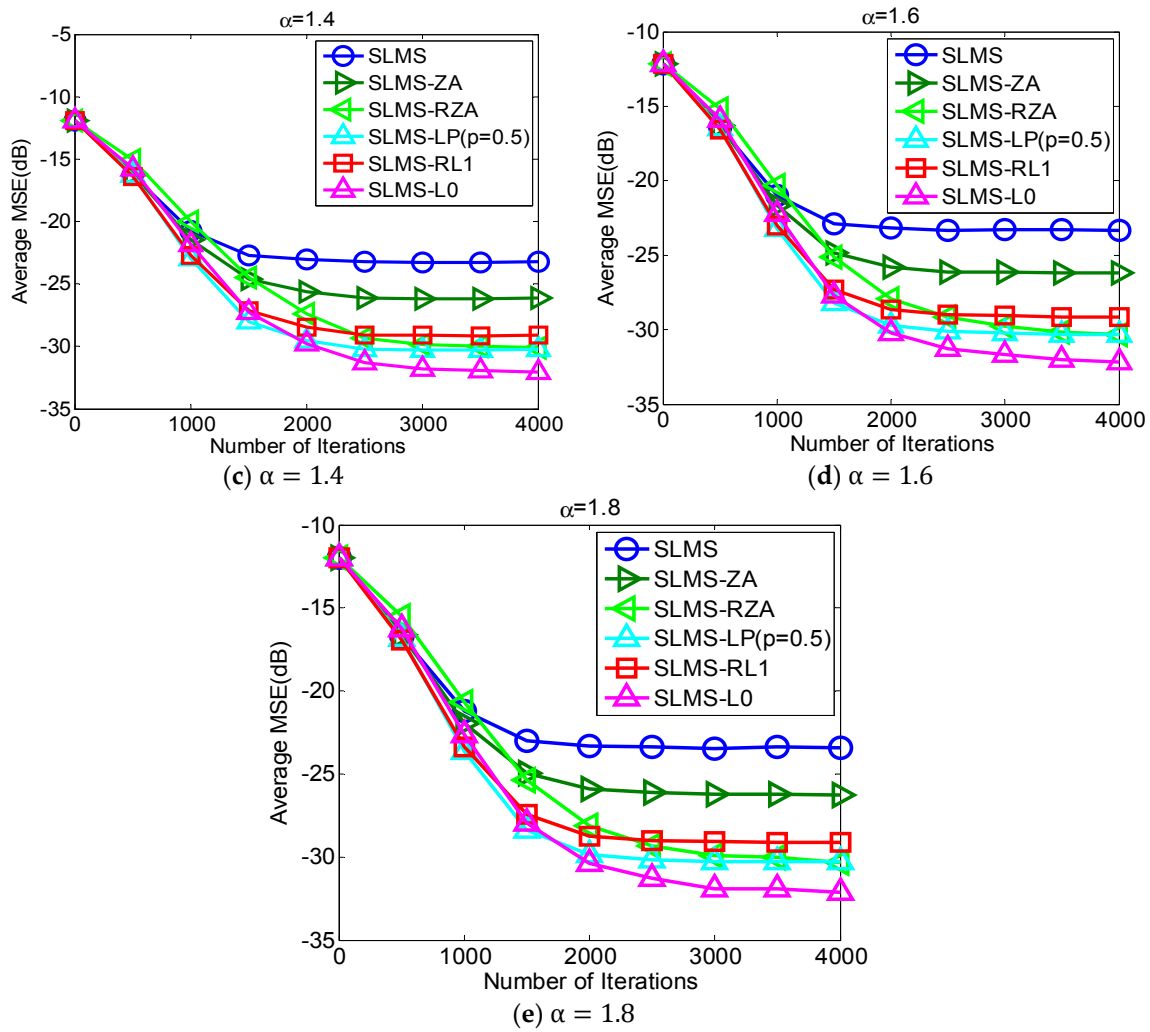
where  $C(n) = E \{v(n)v^T(n)\}$  denotes the second-order moment matrix of channel estimation error vector  $v(n)$ .



**Figure 1.** PDF comparisons of  $\alpha$ -stable noise: (a) symmetric distribution; (b) skewed distribution; (c) scale distribution.



**Figure 2.** Cont.



**Figure 2.** Monte Carlo based MSE curves averaging over 1000 runs with respect to different characteristic exponents  $\alpha \in \{1.0, 1.2, 1.4, 1.6, 1.8\}$ , in scenarios of dispersion parameter  $\gamma = 1.0$ , channel sparsity  $K = 8$ , channel length  $N = 128$ , and  $\text{SNR} = 10$  dB.

### 3. Proposed Sparse SLMS Algorithms

By incorporating sparsity-aware function into the cost function of the standard SLMS in Equation (6), sparse SLMS algorithms could be developed to take advantage of sparse structure information, to mitigate S&S noise as well as to reconstruct channel FIR. First of all, this section proposes five effective sparse SLMS algorithms with different sparse constraints. These proposed algorithms are SLMS-ZA, SLMS-RZA, SLMS-RL1, SLMS-LP, and SLMS-L0. Later, performance analysis will be given to confirm the effectiveness of the proposed algorithms.

#### 3.1. First Proposed Algorithm: SLMS-ZA

The cost function of the LMS-ZA algorithm [7] was developed as

$$G_{ZA}(w(n)) = (1/2) e^2(n) + \lambda_{ZA} ||w(n)||_1 \quad (17)$$

where  $\lambda_{ZA}$  stands for a positive parameter to trade off instantaneous estimation square error and sparse penalty of  $w(n)$ . It is worth noting that the optimal selection of  $\lambda_{ZA}$  is very difficult due to the fact that  $\lambda_{ZA}$  depends on many variables such as channel sparsity, instantaneous updating error  $e(n)$ ,

SNR, and so on. Throughout this paper, the regularization parameter will be selected empirically via the Monte Carlo method. According to Equation (17), the LMS-ZA algorithm was developed as

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{w}(n) - \mu \frac{\partial G_{ZA}(\mathbf{w}(n))}{\partial \mathbf{w}(n)} \\ &= \mathbf{w}(n) + \mu e(n) \mathbf{x}(n) - \rho_{ZA} \text{sgn}(\mathbf{w}(n)) \end{aligned} \quad (18)$$

where  $\rho_{ZA} = \mu \cdot \lambda_{ZA}$  depends the  $\mu$  and  $\lambda_{ZA}$ . To mitigate the S&S noises, by constraining sign function on  $e(n)$ , the SLMS-ZA algorithm is developed as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \text{sgn}(e(n)) \mathbf{x}(n) - \rho_{ZA} \text{sgn}(\mathbf{w}(n)) \quad (19)$$

where the first  $\text{sgn}(\cdot)$  function is utilized to remove impulsive noise in  $e(n)$  while the second one acts as a sparsity-inducing function to exploit channel sparsity in  $\mathbf{w}(n)$ . Please note that the steady-state mean square error (MSE) performance of the proposed SLMS-ZA depends highly on  $\rho_{ZA}$ .

### 3.2. Second Proposed Algorithm: SLMS-RZA

A stronger sparse penalty function can obtain more accurate sparse information [19]. By devising an improved sparse penalty function RZA, we can develop an improved LMS-RZA algorithm. Its cost function can be constructed as

$$G_{RZA}(w_i(n)) = \frac{1}{2} e^2(n) + \lambda_{RZA} \sum_{i=0}^{N-1} \log(1 + \varepsilon_{RZA} |w_i(n)|) \quad (20)$$

where  $\lambda_{RZA} > 0$  is a positive parameter. By deriving Equation (20), the update equation of SLMS-RZA is obtained as

$$\begin{aligned} w_i(n+1) &= w_i(n) + \mu \frac{\partial G_{RZA}(w_i(n))}{\partial w_i(n)} \\ &= w_i(n) + \mu e(n) x(n-i) - \frac{\rho_{RZA} \text{sgn}(w_i(n))}{1 + \varepsilon_{RZA} |w_i(n)|} \end{aligned} \quad (21)$$

where  $\rho_{RZA} = \mu \lambda_{RZA} \varepsilon_{RZA}$ . By collecting all of the coefficients as the matrix-vector form, Equation (21) can be expressed as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n) \mathbf{x}(n) - \frac{\rho_{RZA} \text{sgn}(\mathbf{w}(n))}{1 + \varepsilon_{RZA} |\mathbf{w}(n)|} \quad (22)$$

where  $\varepsilon_{RZA} = 20$  [7] denotes reweighted factor. By inducing sign function to constraint  $e(n)$  in Equation (22), the stable SLMS-RZA algorithm is proposed as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \text{sgn}(e(n)) \mathbf{x}(n) - \frac{\rho_{RZA} \text{sgn}(\mathbf{w}(n))}{1 + \varepsilon_{RZA} |\mathbf{w}(n)|} \quad (23)$$

### 3.3. Third Proposed Algorithm: SLMS-RL1

In addition to the RZA in (23), RL1 function was also considered as an effective sparse constraint in the field of compressive sensing (CS) [19]. By choosing a suitable reweighted factor,  $\delta_{RL1}$ , RL1 could approach the optimal  $\ell_0$ -norm (L0) constraint. Hence, the LMS-RL1 algorithm has been considered an attractive application of sparse channel estimation. The LMS-RL1 algorithm [5] was developed as

$$G_{RL1}(\mathbf{w}(n)) = (1/2) e^2(n) + \lambda_{RL1} ||\mathbf{f}(n) \mathbf{w}(n)||_1 \quad (24)$$

where  $\lambda_{RL1}$  denotes positive regularization parameter and  $\mathbf{f}(n)$  is defined as

$$[\mathbf{f}(n)]_i = \frac{1}{\delta_{RL1} + |[\mathbf{w}(n-1)]_i|}, \quad i = 0, 1, \dots, N-1 \quad (25)$$

where  $\delta_{RL1} > 0$  and then  $[f(n)]_i > 0$ . By taking the derivation of Equation (24), the LMS-RL1 algorithm was updated as

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{w}(n) + \mu \frac{\partial G_{RL1}(\mathbf{w}(n))}{\partial \mathbf{w}(n)} \\ &= \mathbf{w}(n) + \mu e(n) \mathbf{x}(n) - \rho_{RL1} \text{sgn}(\mathbf{f}(n) \mathbf{w}(n)) \mathbf{f}^T(n) \\ &= \mathbf{w}(n) + \mu e(n) \mathbf{x}(n) - \rho_{RL1} \text{sgn}(\mathbf{w}(n)) \mathbf{f}^T(n) \\ &= \mathbf{w}(n) + \mu e(n) \mathbf{x}(n) - \frac{\rho_{RL1} \text{sgn}(\mathbf{w}(n))}{\delta_{RL1} + |\mathbf{w}(n-1)|} \end{aligned} \quad (26)$$

where  $\rho_{RL1} = \mu \lambda_{RL1}$ . The third step of derivation can be obtained since  $\text{sgn}(\mathbf{f}(n)) = \mathbf{1}_{1 \times N}$  and then  $\text{sgn}(\mathbf{f}(n) \mathbf{w}(n)) = \text{sgn}(\mathbf{w}(n))$ . Here the cost function  $G_{RL1}(n)$  is convex due to the fact that it does not depend on  $\mathbf{w}(n)$ . Likewise, sign function is adopted for constraint  $e(n)$  and then robust SLMS-RL1 algorithm is obtained as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \text{sgn}(e(n)) \mathbf{x}(n) - \frac{\rho_{RL1} \text{sgn}(\mathbf{w}(n))}{\delta_{RL1} + |\mathbf{w}(n-1)|} \quad (27)$$

### 3.4. Fourth Proposed Algorithm: SLMS-LP

$\ell_p$ -norm sparse penalty is a nonconvex function to exploit sparse prior information. In [5], LMS-LP based channel estimation algorithms was developed as

$$G_{LP}(\mathbf{w}(n)) = (1/2) e^2(n) + \lambda_{LP} ||\mathbf{w}(n)||_p \quad (28)$$

where  $\lambda_{LP} > 0$  is a positive parameter. The update function of LMS-LP is given as

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{w}(n) - \mu \frac{\partial G_{LP}(\mathbf{w}(n))}{\partial \mathbf{w}(n)} \\ &= \mathbf{w}(n) + \mu e(n) \mathbf{x}(n) - \frac{\rho_{LP} ||\mathbf{w}(n)||_p^{1-p} \text{sgn}(\mathbf{w}(n))}{\varepsilon_{LP} + |\mathbf{w}(n)|^{1-p}} \end{aligned} \quad (29)$$

where  $\varepsilon_{LP} > 0$  is a threshold parameter and  $\rho_{LP} = \mu \lambda_{LP}$ . To remove the S&S noise, SLMS-LP based robust adaptive channel estimation is written as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \text{sgn}(e(n)) \mathbf{x}(n) - \frac{\rho_{LP} ||\mathbf{w}(n)||_p^{1-p} \text{sgn}(\mathbf{w}(n))}{\varepsilon_{LP} + |\mathbf{w}(n)|^{1-p}} \quad (30)$$

### 3.5. Fifth Proposed Algorithm: SLMS-L0

It is well known that the  $\ell_0$ -norm penalty can exploit the sparse structure information. Hence, the L0-LMS algorithm is constructed as

$$G_{L0}(\mathbf{w}(n)) = \frac{1}{2} e^2(n) + \lambda_{L0} ||\mathbf{w}(n)||_0 \quad (31)$$

where  $\lambda_{L0} > 0$  and  $||\mathbf{w}(n)||_0$  stands for optimal  $\ell_0$ -norm function. However, it is a NP-hard problem to solve the  $\ell_0$ -norm sparse minimization [22]. The NP-hard problem in Equation (31) can be solved by an approximate continuous function:

$$||\mathbf{w}(n)||_0 \approx \sum_{i=0}^{N-1} \left(1 - e^{-\theta |w_i(n)|}\right) \quad (32)$$

Then, the previous cost function (29) is changed to

$$G_{L0}(\mathbf{w}(n)) = \frac{1}{2} e^2(n) + \lambda_{L0} \sum_{i=0}^{N-1} \left(1 - e^{-\theta |w_i(n)|}\right) \quad (33)$$

The first-order Taylor series expansion of exponential function  $e^{-\theta|w_i(n)|}$  can be expressed as

$$e^{-\theta|w_i(n)|} \approx \begin{cases} 1 - \theta |w_i(n)|, & \text{when } |w_i(n)| \leq 1/\theta \\ 0, & \text{others} \end{cases} \quad (34)$$

Then, the LMS-L0 based robust adaptive sparse channel estimation algorithm is given as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n) - \rho_{L0} \text{sgn}(\mathbf{w}(n)) e^{-\theta|\mathbf{w}(n)|} \quad (35)$$

where  $\rho_{L0} = \mu\lambda_{L0}$ . In Equation (35), the exponential function still exhausts high computational resources. To further reduce it, a simple approximation function  $\zeta(\mathbf{w}(n))$  is proposed in [20]. By introducing sign function into Equation (35), the SLMS-L0 based robust channel estimation algorithm is written as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \text{sgn}(e(n)) \mathbf{x}(n) - \rho_{L0} \zeta(\mathbf{w}(n)) \quad (36)$$

where  $\rho_{L0} = \mu\lambda_{L0}$  and  $\zeta(w_i(n))$  are defined as

$$\zeta(w_i(n)) = \begin{cases} 2\theta^2 w_i(n) - 2\theta \text{sgn}(w_i(n)), & \text{when } |w_i(n)| \leq 1/\theta \\ 0, & \text{others} \end{cases} \quad (37)$$

for  $\zeta(\mathbf{w}(n)) = [\zeta(w_0(n)), \zeta(w_1(n)), \dots, \zeta(w_{N-1}(n))]^T$ .

#### 4. Convergence Analysis of the Proposed Algorithms

Unlike the standard SLMS algorithm, the proposed sparse SLMS algorithms can further improve estimation accuracy by exploiting channel sparsity. For convenience of theoretical analysis without loss of generality, the above proposed sparse SLMS algorithms are generalized as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \text{sgn}(e(n)) \mathbf{x}(n) - \rho f(\mathbf{w}(n)) \quad (38)$$

where  $f(\mathbf{w}(n))$  denotes sparsity constraint function and  $\rho$  denotes the regularization parameter. Throughout this paper, our analysis is based on independent assumptions as below:

$$E(z(n)\mathbf{x}(n)) = 0 \quad (39)$$

$$E[z(n)\mathbf{x}(n)\mathbf{v}^T(n)] = 0 \quad (40)$$

$$E[z(n)\mathbf{v}(n)\mathbf{x}^T(n)] = 0 \quad (41)$$

$$E[z(n)f(\mathbf{w}(n))\mathbf{x}^T(n)] = 0 \quad (42)$$

$$E[z(n)\mathbf{x}(n)f^T(\mathbf{w}(n))] = 0 \quad (43)$$

**Theorem 1.** If  $\mu$  satisfies (15), the mean coefficient vector  $E\{\mathbf{w}(n)\}$  approaches

$$\mathbf{w}(\infty) = \mathbf{w} - \sqrt{\pi/2}\mu^{-1}\rho\mathbf{R}^{-1}\gamma^{1/\alpha}f(\mathbf{w}(\infty)) \quad (44)$$

**Proof.** By subtracting  $\mathbf{w}$  from both sides of (35), the mean estimation error  $E\{\mathbf{v}(n+1)\}$  is derived as

$$\begin{aligned}
 E\{v(n+1)\} &= E\{v(n)\} + \mu E\{\text{sgn}(e(n))x(n)\} - \rho f(w(n)) \\
 &\approx E\{v(n)\} + \mu E\{\text{sgn}(e(n))x(n) | v(n)\} - \rho f(w(n)) \\
 &= E\{v(n)\} + \sqrt{2/\pi} \mu \sigma_e^{-1}(n) E\{e(n)x(n)\} - \rho f(w(n)) \\
 &= \{I - \sqrt{2/\pi} \mu \sigma_e^{-1}(n) R\} E\{v(n)\} - \rho f(w(n))
 \end{aligned} \tag{45}$$

It is worth mentioning that vector  $\rho f(w(n))$  is bounded for all sparse constraints. For example, if  $f(w(n)) = \text{sgn}(w(n))$ , then the bound is between  $-\rho I_N$  and  $\rho I_N$ , where  $I_N$  is an  $N$ -length identity vector. For  $n \rightarrow \infty$ , Equation (45) can be rewritten as

$$\begin{aligned}
 E\{v(\infty)\} &= \{I - \sqrt{2/\pi} \mu \sigma_e^{-1}(\infty) R\} E\{v(\infty)\} - \rho f(w(\infty)) \\
 &\approx \{I - \sqrt{2/\pi} \mu \gamma^{-1/\alpha} R\} E\{v(\infty)\} - \rho f(w(\infty))
 \end{aligned} \tag{46}$$

where  $\text{Tr}(RC(\infty)) = \sigma_x^2 c(\infty) \ll \gamma^{1/\alpha}$  and

$$\sigma_e^2(\infty) = \lim_{n \rightarrow \infty} \sigma_e^2(n) = \gamma^{2/\alpha} + \text{Tr}(RC(\infty)) \approx \gamma^{2/\alpha} \tag{47}$$

are utilized in the above equation. Since  $E\{w(\infty)\} = w - E\{v(\infty)\}$ , according to Equation (47), one can easily get Theorem 1.

**Theorem 2.** Let  $\Omega$  denotes the index set of nonzero taps, i.e.,  $w_i \neq 0$  for  $i \in \Omega$ . Assuming  $\rho$  is sufficiently small so that for every  $i \in \Omega$ , the excess MSE of sparse SLMS algorithms is

$$P_{ex}(\infty) = \frac{\mu \gamma^{1/\alpha} \varphi_2}{\sqrt{2\pi} \varphi_1} + \frac{\rho \gamma^{1/\alpha} \eta'_1}{\sqrt{2\pi} \mu \varphi_1} \left( \rho - \frac{\eta'_2}{\eta'_1} \right) \tag{48}$$

where  $\varphi_1$ ,  $\varphi_2$ ,  $\eta'_1$  and  $\eta'_2$  are defined as:

$$\varphi_1 = \text{Tr}[(I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1}] \tag{49}$$

$$\varphi_2 = \text{Tr}[R(I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1}] \triangleleft \tag{50}$$

$$\eta'_1 \triangleq \frac{N}{1 - \sqrt{2/\pi} \mu \gamma^{-1/\alpha} \lambda_{\max}} \tag{51}$$

$$\eta'_2 \triangleq 2E\left\{\text{Tr}\left[w(\infty)f^T(w(\infty))\right]\right\} - 2E\left\{\text{Tr}\left[w f^T(w(\infty))\right]\right\} \tag{52}$$

**Proof.** By using the above independent assumptions in Equations (49)–(52), the second moment  $C(n+1)$  of the weight error vector  $v(n+1)$  can be evaluated recursively as

$$\begin{aligned}
 C(n+1) &= E\{v(n+1)v^T(n+1)\} \\
 &= E\left\{[v(n) + \mu \text{sgn}(e(n))x(n) - \rho f(w(n))][v(n) + \mu \text{sgn}(e(n))x(n) - \rho f(w(n))]^T\right\} \\
 &= E\{v(n)v^T(n)\} + \mu^2 E\{x(n)x^T(n)\} \\
 &\quad + \mu \left\{ E[\text{sgn}(e(n))v(n)x^T(n)] + E[\text{sgn}(e(n))x(n)v^T(n)] \right\} \\
 &\quad - \underbrace{\mu \rho \left\{ E[\text{sgn}(e(n))x(n)f^T(w(n))] + E[\text{sgn}(e(n))f(w(n))x^T(n)] \right\}}_{A_1(n)} \\
 &\quad - \underbrace{\rho \left\{ E[v(n)f^T(w(n))] + E[f(w(n))v^T(n)] \right\}}_{A_2(n)} + \underbrace{\rho^2 E\{f(w(n))f^T(w(n))\}}_{A_4(n)} \\
 &= C(n) + \underbrace{\mu^2 R + A_1(n) - (A_2(n) + A_3(n))}_{A_5(n)} + A_4(n)
 \end{aligned} \tag{53}$$

where  $A_1(n)$ ,  $A_2(n)$ , and  $A_5(n)$  are further derived as

$$\begin{aligned} A_1(n) &\approx \mu E \{ \text{sgn}(e(n)) v(n) x^T(n) | v(n) \} + \mu E \{ \text{sgn}(e(n)) x(n) v^T(n) | v(n) \} \\ &= -\sqrt{\pi/2} \mu \sigma_e^{-1}(n) \{ E [v(n) E [e(n) x^T(n) | v(n)]] + E [E [e(n) x(n) | v(n)] v^T(n)] \} \\ &= -\sqrt{\pi/2} \mu \sigma_e^{-1}(n) \{ E [v(n) v^T(n)] R + R E [v(n) v^T(n)] \} \\ &= -\sqrt{\pi/2} \mu \sigma_e^{-1}(n) [C(n)R + RC(n)] \end{aligned} \quad (54)$$

$$\begin{aligned} A_2(n) &= \mu \rho \{ E [\text{sgn}(e(n)) x(n) f^T(w(n))] + E [\text{sgn}(e(n)) f(w(n)) x^T(n)] \} \\ &\approx \mu \rho E \{ E [\text{sgn}(e(n)) x(n) f^T(w(n)) | v(n)] + E [\text{sgn}(e(n)) f(w(n)) x^T(n) | v(n)] \} \\ &= -\sqrt{2/\pi} \mu \rho \sigma_e^{-1} E \{ E [x(n) e(n) | v(n)] f^T(w(n)) + f(w(n)) E [e(n) x^T(n) | v(n)] \} \\ &= -\sqrt{2/\pi} \mu \rho \sigma_e^{-1} \{ R E [v(n) f^T(w(n))] + E [f(w(n)) v^T(n)] R \} \end{aligned} \quad (55)$$

$$\begin{aligned} A_5(n) &\triangleq A_2(n) + A_3(n) \\ &= -\sqrt{2/\pi} \mu \rho \sigma_e^{-1} \{ R E [v(n) f^T(w(n))] + E [f(w(n)) v^T(n)] R \} \\ &\quad + \rho \{ E [v(n) f^T(w(n))] + E [f(w(n)) v^T(n)] \} \\ &= \rho (I - \sqrt{2/\pi} \mu \sigma_e^{-1} R) E [v(n) f^T(w(n))] + \rho E [f(w(n)) v^T(n)] (I - \sqrt{2/\pi} \mu \sigma_e^{-1} R) \end{aligned} \quad (56)$$

By substituting Equations (54) and (56) into Equation (53), we obtain

$$\begin{aligned} C(n+1) &= C(n) + \mu^2 R + A_1(n) - A_5(n) + A_4(n) \\ &= C(n) + \mu^2 R - \sqrt{2/\pi} \mu \sigma_e^{-1}(n) [C(n)R + RC(n)] - A_5(n) + A_4(n) \end{aligned} \quad (57)$$

Letting  $n \rightarrow \infty$  and using Equation (47), Equation (57) is further rewritten as

$$C(\infty)R + RC(\infty) \approx \sqrt{2/\pi} \mu \gamma^{1/\alpha} R + \sqrt{2/\pi} \mu^{-1} \gamma^{1/\alpha} \lim_{n \rightarrow \infty} [A_4(n) - A_5(n)] \quad (58)$$

Multiplying both sides of Equation (58) by  $(I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1}$  from right, the following can be derived as

$$\begin{aligned} [C(\infty)R + RC(\infty)] (I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1} &\approx \sqrt{2/\pi} \mu \gamma^{1/\alpha} R (I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1} \\ &\quad + \sqrt{2/\pi} \mu^{-1} \gamma^{1/\alpha} \lim_{n \rightarrow \infty} A_4(n) (I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1} \\ &\quad - \sqrt{2/\pi} \mu^{-1} \gamma^{1/\alpha} \lim_{n \rightarrow \infty} A_5(n) (I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1} \end{aligned} \quad (59)$$

Taking the trace of the two sides of Equation (59), since  $\text{Tr}(C(\infty)R) = \text{Tr}(RC(\infty))$ , the excess MSE is derived as

$$\begin{aligned} P_{ex}(\infty) &= \text{Tr}[RC(\infty)] \\ &= \underbrace{\frac{\sqrt{2/\pi} \mu \gamma^{1/\alpha}}{2\varphi_1} \text{Tr} \left[ R (I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1} \right]}_{\varphi_2} \\ &\quad + \underbrace{\frac{\sqrt{2/\pi} \mu^{-1} \gamma^{1/\alpha}}{2\varphi_1} \lim_{n \rightarrow \infty} \text{Tr} \left[ A_4(n) (I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1} \right]}_{\eta_1(n)} \\ &\quad - \underbrace{\frac{\sqrt{2/\pi} \mu^{-1} \gamma^{1/\alpha}}{2\varphi_1} \lim_{n \rightarrow \infty} \text{Tr} \left[ A_5(n) (I - \sqrt{2/\pi} \mu \sigma_e^{-1} R)^{-1} \right]}_{\eta_2(n)} \end{aligned} \quad (60)$$

The matrix  $(I - \sqrt{2/\pi} \mu \gamma^{-1/\alpha} R)$  is symmetric, and its eigenvalue decomposition can be written as

$$I - \sqrt{2/\pi} \mu \gamma^{-1/\alpha} R = UVU^T \quad (61)$$

with  $\mathbf{U}$  being the orthonormal matrix of eigenvectors and  $\mathbf{V}$  being a diagonal matrix of eigenvalues. Therefore,  $(\mathbf{I} - \sqrt{2/\pi\mu}\sigma_e^{-1}\mathbf{R}) = \mathbf{U}\mathbf{V}^{-1}\mathbf{U}^T$ . Let  $\lambda_{\max}$  be the largest eigenvalue of the covariance matrix  $\mathbf{R}$  and  $\mu$  be small enough such that  $(1 - \sqrt{2/\pi\mu}\gamma^{-1/\alpha}\lambda_{\max})^{-1} > 0$  under certain noise cases. Since  $\mathbf{V}^{-1}$  is a diagonal matrix, elements are all non-negative and less than or equal to  $(1 - \sqrt{2/\pi\mu}\gamma^{-1/\alpha}\lambda_{\max})^{-1}$ , hence,  $\eta_1 \triangleq \lim_{n \rightarrow \infty} \eta_1(n)$  and  $\eta_2 \triangleq \lim_{n \rightarrow \infty} \eta_2(n)$  are further derived as

$$\begin{aligned}\eta_1 &\triangleq \lim_{n \rightarrow \infty} \eta(n) = \lim_{n \rightarrow \infty} \text{Tr} \left\{ \mathbf{A}_4 (\mathbf{I} - \sqrt{2/\pi\mu}\sigma_e^{-1}\mathbf{R})^{-1} \right\} \\ &= \lim_{n \rightarrow \infty} \rho^2 E \left\{ \text{Tr} \left[ f(\mathbf{w}(n)) (\mathbf{I} - \sqrt{2/\pi\mu}\gamma^{1/\alpha}\mathbf{R})^{-1} f^T(\mathbf{w}(n)) \right] \right\} \\ &= \lim_{n \rightarrow \infty} \rho^2 E \left\{ \text{Tr} [f(\mathbf{w}(n)) \mathbf{U}\mathbf{V}^{-1}\mathbf{U}^T f^T(\mathbf{w}(n))] \right\} \\ &= \lim_{n \rightarrow \infty} \rho^2 E \left\{ \text{Tr} [\mathbf{V}^{-1}\mathbf{U}^T f(\mathbf{w}(n)) f^T(\mathbf{w}(n)) \mathbf{U}] \right\} \\ &\leq \lim_{n \rightarrow \infty} \frac{\rho^2}{1 - \sqrt{2/\pi\mu}\gamma^{-1/\alpha}\lambda_{\max}} E \left\{ \text{Tr} [\mathbf{U}^T f(\mathbf{w}(n)) f^T(\mathbf{w}(n)) \mathbf{U}] \right\} \\ &= \frac{\rho^2}{1 - \sqrt{2/\pi\mu}\gamma^{-1/\alpha}\lambda_{\max}} E \left\{ \text{Tr} [f^T(\mathbf{w}(\infty)) \mathbf{U}\mathbf{U}^T f(\mathbf{w}(\infty))] \right\} \\ &\leq \frac{\rho^2}{1 - \sqrt{2/\pi\mu}\gamma^{-1/\alpha}\lambda_{\max}} E \left\{ f^T(\mathbf{w}(\infty)) f(\mathbf{w}(\infty)) \right\} \leq \frac{N\rho^2}{1 - \sqrt{2/\pi\mu}\gamma^{-1/\alpha}\lambda_{\max}}\end{aligned}\quad (62)$$

$$\begin{aligned}\eta_2 &\triangleq \lim_{n \rightarrow \infty} \eta_2(n) = \lim_{n \rightarrow \infty} \text{Tr} \left\{ \mathbf{A}_5 (\mathbf{I} - \sqrt{2/\pi\mu}\sigma_e^{-1}\mathbf{R})^{-1} \right\} \\ &= \lim_{n \rightarrow \infty} \rho E \left\{ \text{Tr} [\mathbf{v}(n) f^T(\mathbf{w}(n)) + f(\mathbf{w}(n)) \mathbf{v}^T(n)] \right\} \\ &= \lim_{n \rightarrow \infty} 2\rho E \left\{ \text{Tr} [\mathbf{v}(n) f^T(\mathbf{w}(n))] \right\} \\ &= \lim_{n \rightarrow \infty} 2\rho E \left\{ \text{Tr} [\mathbf{w}(n) f^T(\mathbf{w}(n))] \right\} - 2\rho E \left\{ \text{Tr} [\mathbf{w} f^T(\mathbf{w}(n))] \right\} \\ &= 2\rho E \left\{ \text{Tr} [\mathbf{w}(\infty) f^T(\mathbf{w}(\infty))] \right\} - 2\rho E \left\{ \text{Tr} [\mathbf{w} f^T(\mathbf{w}(\infty))] \right\}\end{aligned}\quad (63)$$

Substituting Equations (62) and (63) into Equation (60), the excess MSE is finally derived as

$$P_{ex}(\infty) = \frac{\mu\gamma^{1/\alpha}\varphi_2}{\sqrt{2\pi}\varphi_1} + \frac{\rho\gamma^{1/\alpha}\eta'_1}{\sqrt{2\pi}\mu\varphi_1} \left( \rho - \frac{\eta'_2}{\eta'_1} \right) \quad (64)$$

where  $\eta'_1 \triangleq \eta_1/\rho^2$  and  $\eta'_2 \triangleq \eta_2/\rho^2$ . According to Equation (62), one can find that  $\eta'_1$  is bound as

$$0 < \eta'_2 \leq \frac{N}{1 - \sqrt{2/\pi\mu}\gamma^{-1/\alpha}\lambda_{\max}} \quad (65)$$

The excess MSE of sparse SLMS in Equation (65) implies that choosing suitable  $\rho < \eta'_1/\eta'_2$  can lead to smaller excess MSE than standard SLMS algorithm.

## 5. Numerical Simulations and Discussion

To evaluate the proposed robust channel estimation algorithms, we compare these algorithms in terms of channel sparsity and non-Gaussian noise level. A typical broadband wireless communication system is considered in computer simulations [3]. The baseband bandwidth is assigned as 60 MHz and the carrier frequency is set as 2.1 GHz. Signal multipath transmission causes the 1.06  $\mu\text{s}$  delay spread. According to the Shannon sampling theory, the channel length is equivalent to  $n = 128$ . In addition, average mean square error (MSE) is adopted for evaluate the estimation error. The MSE is defined as

$$\text{MSE} \{ \mathbf{w}(n) \} (\text{dB}) \triangleq 10\log_{10} (1/M) \sum_{m=1}^M \|\mathbf{w}(n) - \mathbf{w}\|_2^2 / \|\mathbf{w}\|_2^2 \quad (66)$$

where  $w$  and  $\hat{w}$  denote actual channel and estimator, respectively.  $M = 1000$  independent runs are adopted for Monte Carlo simulations. The nonzero channel taps are generated to satisfy random Gaussian distribution as  $\mathcal{CN}(0, \sigma_w^2 I)$  and all of these positions are randomly allocated within  $w$ , which is normalized as  $E\{\|w\|_2^2\} = 1$ . All of the simulation parameters are listed in Table 1.

**Table 1.** List of computer simulation parameters in robust adaptive sparse channel estimation.

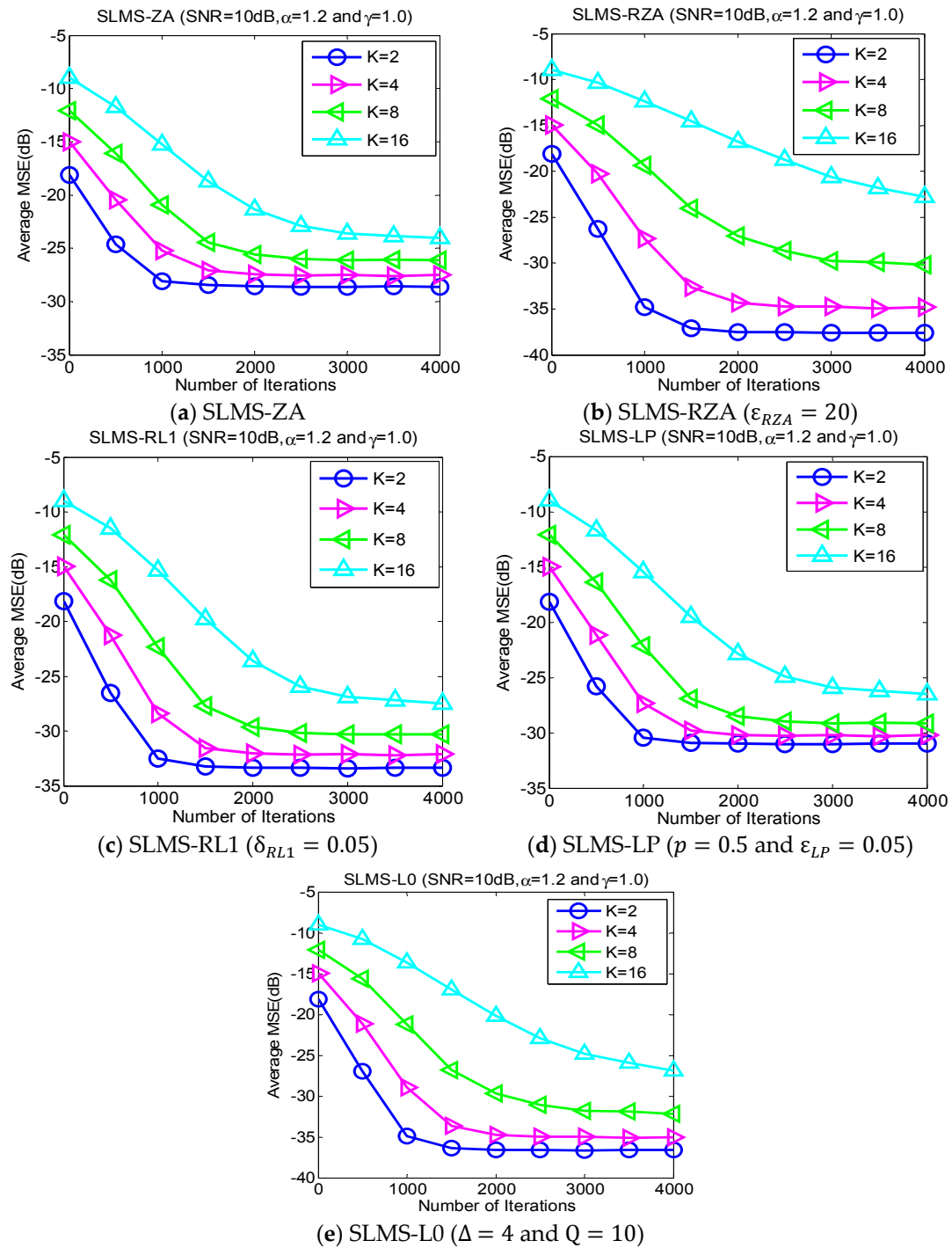
| Parameters                                                                   | Values                                                                                                                                                       |
|------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Training sequence                                                            | Pseudo-random binary sequence                                                                                                                                |
| Non-Gaussian noise level                                                     | $\alpha \in \{1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$ ,<br>$\beta = 0, \gamma \in \{0.5, 1.0, 1.5, 2.0, 2.5\}, \delta = 0$                                           |
| Channel length                                                               | $N = 128$                                                                                                                                                    |
| No. of nonzero coefficients                                                  | $K \in \{2, 4, 8, 16\}$                                                                                                                                      |
| Distribution of nonzero coefficient                                          | Random Gaussian $\mathcal{CN}(0, 1)$                                                                                                                         |
| Received SNR                                                                 | 5 dB $\sim$ 30 dB                                                                                                                                            |
| Gradient descend step size                                                   | $\mu = 0.005$                                                                                                                                                |
| Sparsity-aware positive parameters<br>(i.e., $\rho_{ZA} = \mu\lambda_{ZA}$ ) | $\rho_{ZA} = 2 \times 10^{-4}, \rho_{RZA} = 2 \times 10^{-3}$<br>$\rho_{RL1} = 5 \times 10^{-5}, \rho_{LP} = 5 \times 10^{-6}, \rho_{LP} = 2 \times 10^{-4}$ |
| Reweight parameter of (S)LMS-RZA                                             | $\varepsilon_{RZA} = 20$                                                                                                                                     |
| Threshold parameter of (S)LMS-RL1                                            | $\delta_{RL1} = 0.05$                                                                                                                                        |
| Threshold parameter of (S)LMS-LP                                             | $\varepsilon_{LP} = 0.05$                                                                                                                                    |
| Approximate parameters of (S)LMS-L0                                          | $\Delta = 4, Q = 10$                                                                                                                                         |

### 5.1. Experiment 1. MSE Curves of Proposed Algorithms vs. Different Alpha-Stable Noise

The proposed robust adaptive sparse channel estimation algorithms are evaluated with respect to  $\alpha$  in the scenarios of  $K = 8$  and  $SNR = 10$  dB, as shown in Figure 2. Under different alpha-stable noise regimes, i.e.,  $\alpha \in \{1.0, 1.2, 1.4, 1.6, 1.8\}$ , our proposed algorithms can achieve much better MSE performance than standard SLMS. With different sparsity constraint functions, i.e., ZA, RZA, LP, RL1, and L0, different performance gain could be obtained. Since L0-norm constraint exploits channel sparsity most efficiently in these sparse constraint functions, Figure 2 shows that the lowest MSE of SLMS-L0 results in the lowest MSE. Indeed, Figure 2 implies that taking more channel sparse structure information can obtain more performance gain. Hence, selecting an efficient sparse constraint function is an important step in devising sparse SLMS algorithms. In addition, it is worth noting that the convergence speed of SLMS-RZA is slightly slower than other sparse SLMS algorithms while its steady-state MSE performance as good as SLMS-LP. According to Figure 2, proposed SLMS algorithms are confirmed by simulation results in different impulsive noise cases.

### 5.2. Experiment 2. MSE Curves of Proposed Algorithms vs. Channel Sparsity

Our proposed channel estimation algorithms are evaluated with channel sparsity  $K \in \{2, 4, 8, 16\}$  in the scenarios of  $\alpha = 1.2, \gamma = 1.0, K = 8$  and  $SNR = 10$  dB, as shown in Figure 3. We can find that the proposed sparse SLMS algorithms depend on channel sparsity  $K$ , i.e., our proposed algorithms obtained correspondingly better MSE performance in a scenario of sparser channels. In other words, exploiting more channel sparsity information could produce more performance gain and vice versa. Hence, the proposed methods are effective to exploit channel sparsity as well as remove non-Gaussian  $\alpha$ -stable noise.

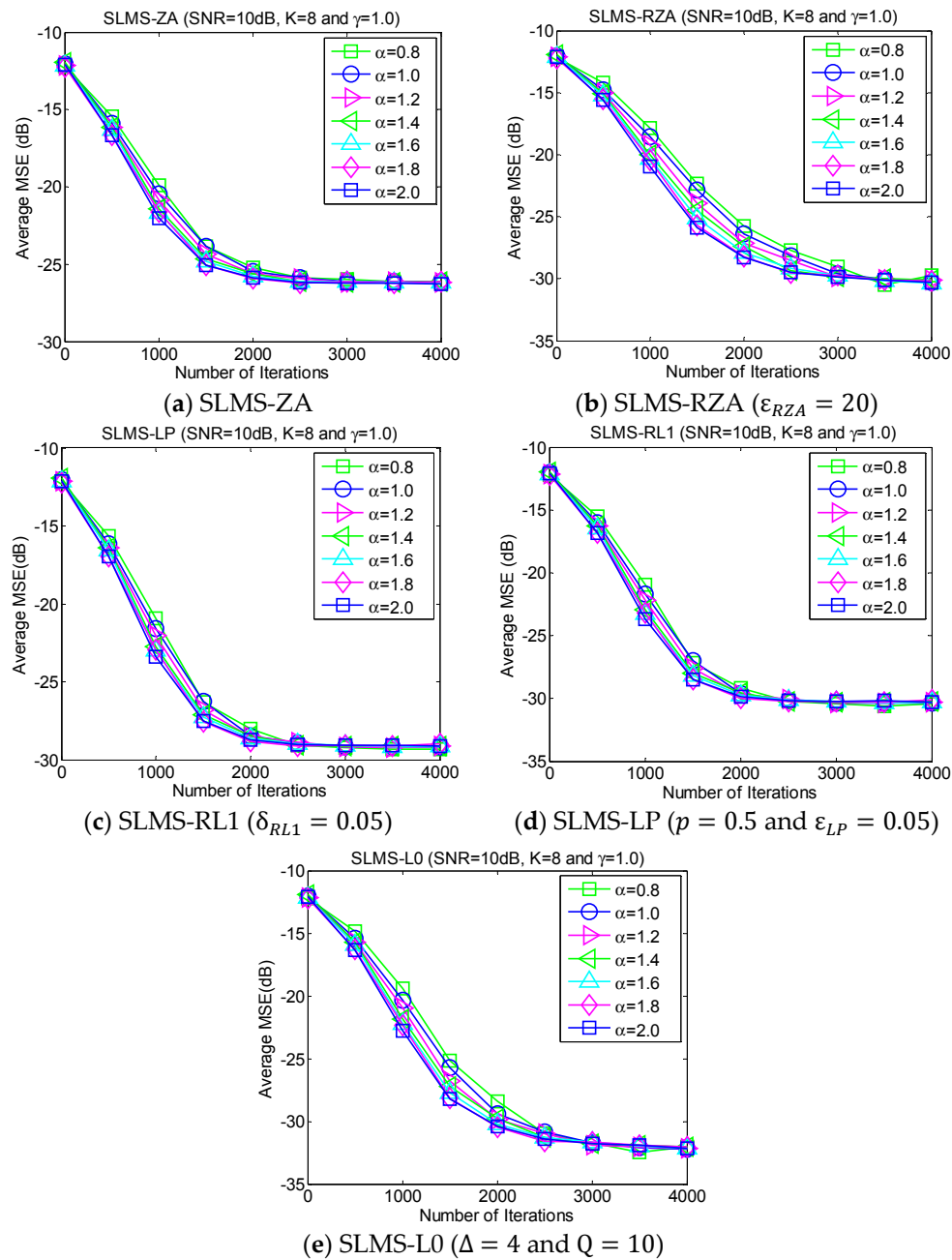


**Figure 3.** Monte Carlo based MSE curves averaging over 1000 runs with respect to different channel sparsities  $K \in \{2, 4, 8, 16\}$ , in scenarios of characteristic exponent  $\alpha = 1.2$ , dispersion parameter  $\gamma = 1.0$ , channel length  $N = 128$ , and SNR = 10 dB.

### 5.3. Experiment 3. MSE Curves of Proposed Algorithms vs. Characteristic Exponent

The average MSE curves of the proposed algorithm with respect to characteristic exponent  $\alpha \in \{0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$  in scenarios of dispersion parameter  $\gamma = 1.0$ , channel length  $N = 128$ , channel sparsity  $K = 8$  and SNR = 10 dB, are depicted as shown in Figure 4. The proposed algorithm is very stable for the different strengths of impulsive noises that are controlled by the characteristic exponent  $\alpha$ . In addition, it is very interesting that the convergence speed of the proposed

SLMS algorithms may be reduced by the relatively small  $\alpha$ . The main reason the proposed SLMS algorithm is utilized is that the sign function is stable at different values of  $\alpha$ . Hence, the proposed algorithm can mitigate non-Gaussian  $\alpha$ -stable noise.

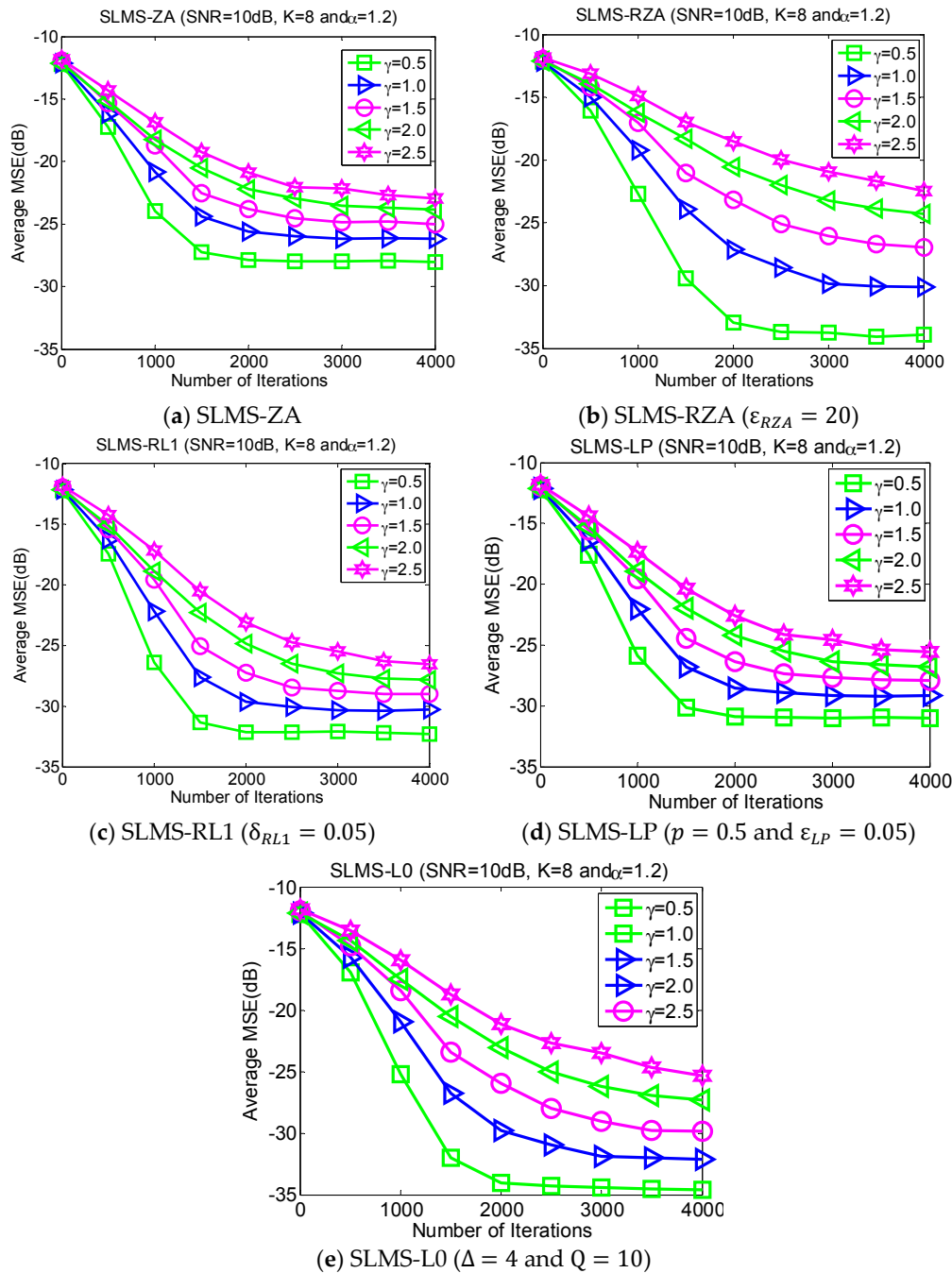


**Figure 4.** Monte Carlo based MSE curves averaging over 1000 runs with respect to different characteristic exponents  $\alpha \in \{0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$ , in scenarios of dispersion parameter  $\gamma = 1.0$ , channel length  $N = 128$ , channel sparsity  $K = 8$ , and SNR = 10 dB.

#### 5.4. Experiment 4. MSE Curves of Proposed Algorithms vs. Dispersive Parameter

Dispersive distribution of  $\alpha$ -stable noise has harmful effects. This experiment evaluates the MSE performance of proposed SLMS algorithms in different dispersive parameters  $\gamma \in \{0.5, 1.0, 1.5, 2.0, 2.5\}$  in the scenarios of  $\alpha = 1.2$ ,  $K = 8$ , and SNR = 10 dB, as shown in Figure 5. Larger  $\gamma$  means more serious dispersion of  $\alpha$ -stable noise and a worse performance for the proposed

algorithm and vice versa. Figure 5 implies that the proposed SLMS algorithms are deteriorated by  $\gamma$  rather than  $\alpha$ . The main reason is that the proposed SLMS algorithm can mitigate the amplitude effect of the  $\alpha$ -stable noise due to the fact that the sign function is utilized in the proposed SLMS algorithms.

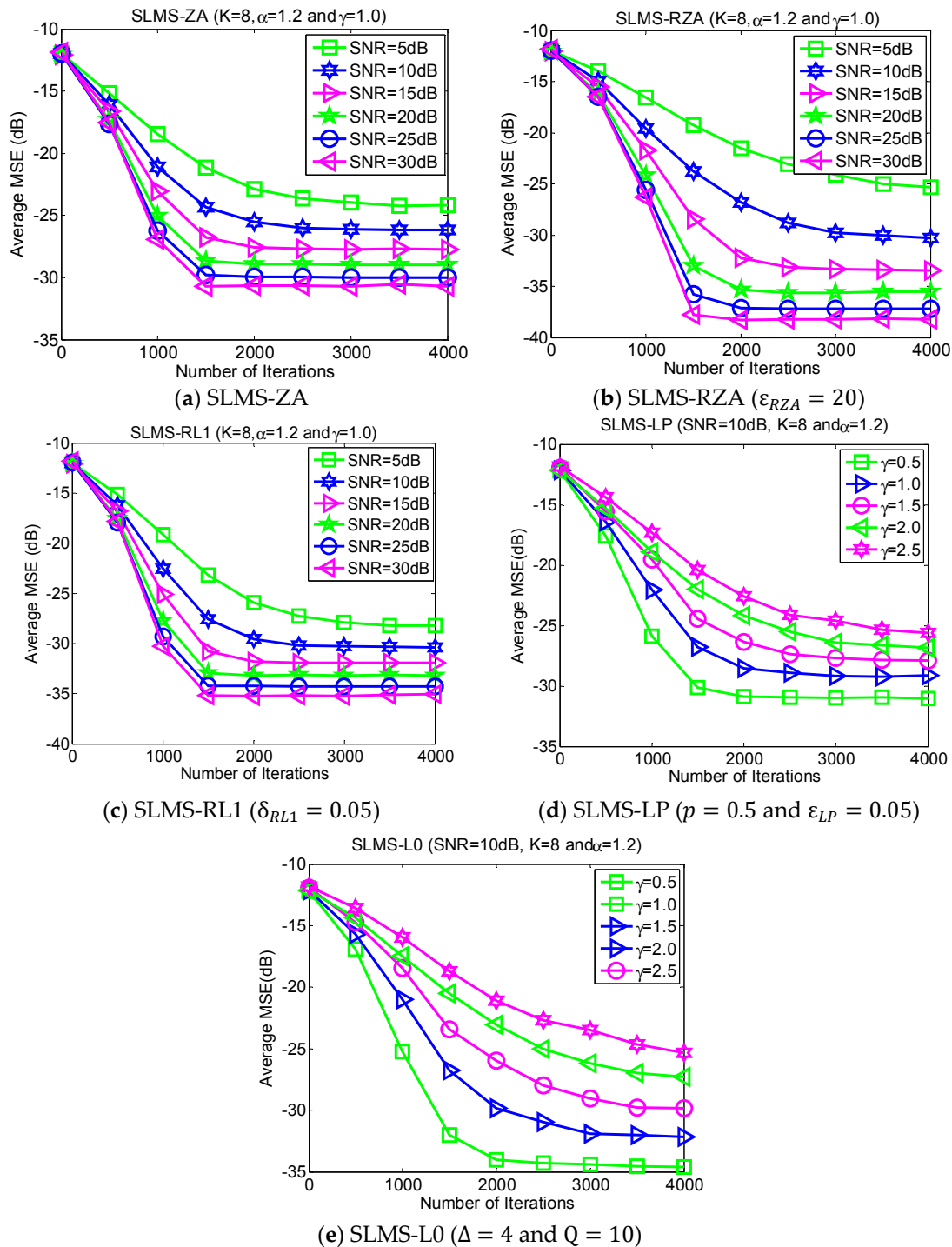


**Figure 5.** Monte Carlo based MSE curves averaging over 1000 runs with respect to different dispersion parameters  $\gamma \in \{0.5, 1.0, 1.5, 2.0, 2.5\}$ , in scenarios of characteristic exponent  $\alpha = 1.2$ , channel length  $N = 128$ , channel sparsity  $K = 8$ , and SNR = 10 dB.

### 5.5. Experiment 5. MSE Curves of Proposed Algorithms vs. SNR

In the different SNR regimes, the average MSE curves of the proposed algorithms are demonstrated in Figure 6 in the scenarios of characteristic exponent  $\alpha = 1.2$ , dispersive parameter

$\gamma = 1.0$ , channel length  $N = 128$ , and channel sparsity  $K = 8$ . The purpose of directing figures in Figure 6 is to further confirm the effectiveness of the proposed algorithms under different SNR regimes.



**Figure 6.** Monte Carlo based MSE curves averaging over 1000 runs with respect to different SNR regimes  $\text{SNR} \in \{5 \text{ dB}, 10 \text{ dB}, 15 \text{ dB}, 20 \text{ dB}, 25 \text{ dB}, 30 \text{ dB}\}$  in scenarios of characteristic exponent  $\alpha = 1.2$ , dispersive parameter  $\gamma = 1.0$ , channel length  $N = 128$ , and channel sparsity  $K = 8$ .

## 6. Conclusions

Based on SxS noise model, we have proposed five sparse SLMS algorithms for robust channel estimation in this paper by introducing sparsity-inducing penalty functions into the standard

SLMS algorithm so that channel sparsity can be exploited to improve the channel estimation performance. Theoretical analysis verified the convergence of the proposed algorithms in terms of mean and excess MSE. Numerical simulations were provided to validate the performance gain of our proposed algorithm.

**Acknowledgments:** The authors would like to express their appreciation to the anonymous reviewers for their constructive comments and positive suggestions. This work was supported by a National Natural Science Foundation of China grant (61271240), the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ1500515), and the project of the Fundamental and Frontier Research plan of Chongqing (cstc2016jcyjA0022).

**Author Contributions:** Tingping Zhang derived the sparse channel estimation algorithms and wrote the paper; Guan Gui and Tingping Zhang performed the experiments; Tingping Zhang wrote the paper; and Guan Gui checked the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Adachi, F.; Garg, D.; Takaoka, S.; Takeda, K. Broadband CDMA techniques. *IEEE Wirel. Commun.* **2005**, *12*, 8–18. [[CrossRef](#)]
- Raychaudhuri, B.D.; Mandayam, N.B. Frontiers of wireless and mobile communications. *Proc. IEEE* **2012**, *100*, 824–840. [[CrossRef](#)]
- Dai, L.; Wang, Z.; Yang, Z. Next-generation digital television terrestrial broadcasting systems: Key technologies and research trends. *IEEE Commun. Mag.* **2012**, *50*, 150–158. [[CrossRef](#)]
- Haykin, S. *Adaptive Filter Theory*; Prentice-Hall: Englewood Cliffs (Upper Saddle River), NJ, USA, 1996.
- Taheri, O.; Vorobyov, S.A. Sparse channel estimation with Lp-norm and reweighted L1-norm penalized least mean squares. In Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, 22–27 May 2011; pp. 2864–2867.
- Gui, G.; Peng, W.; Adachi, F. Improved adaptive sparse channel estimation based on the least mean square algorithm. In Proceedings of the 2013 IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, 7–10 April 2013; pp. 3105–3109.
- Chen, Y.; Gu, Y.; Hero, A.O., III. Sparse LMS for system identification. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 3125–3128.
- Chen, B.; Zhao, S.; Zhu, P.; Principe, J.C. Quantized kernel least mean square algorithm. *IEEE Trans. Neural Netw. Learn. Syst.* **2012**, *23*, 22–32. [[CrossRef](#)] [[PubMed](#)]
- Li, Y.; Hamamura, M. Zero-attracting variable-step-size least mean square algorithms for adaptive sparse channel estimation. *Int. J. Adapt. Control Signal Process.* **2015**. [[CrossRef](#)]
- Li, Y.; Li, W.; Yu, W.; Wan, J.; Li, Z. Sparse adaptive channel estimation based on lp-norm-penalized affine projection algorithm. *Int. J. Antennas Propag.* **2014**. [[CrossRef](#)]
- Lin, J.; Member, S.; Nassar, M.; Evans, B.L. Impulsive noise mitigation in powerline communications using sparse Bayesian learning. *IEEE J. Sel. Areas Commun.* **2013**, *31*, 1172–1183. [[CrossRef](#)]
- Shao, M.; Nikias, C.L. Signal processing with fractional lower order moments: Stable processes and their applications. *Proc. IEEE* **1993**, *81*, 986–1010. [[CrossRef](#)]
- Middleton, D. Non-Gaussian noise models in signal processing for telecommunications: New methods and results for Class A and Class B noise models. *IEEE Trans. Inf. Theory* **1999**, *45*, 1129–1149. [[CrossRef](#)]
- Li, Y.-P.; Lee, T.-S.; Wu, B.-F. A variable step-size sign algorithm for channel estimation. *Signal Process.* **2014**, *102*, 304–312. [[CrossRef](#)]
- Shao, T.; Zheng, Y.; Benesty, J. An affine projection sign algorithm robust against impulsive interferences. *IEEE Signal Process. Lett.* **2012**, *17*, 327–330. [[CrossRef](#)]
- Yoo, J.; Shin, J.; Park, P. Variable step-size affine projection sign algorithm. *IEEE Trans. Circuits Syst. Express Briefs* **2014**, *61*, 274–278.
- Dai, L.; Gui, G.; Wang, Z.; Yang, Z.; Adachi, F. Reliable and energy-efficient OFDM based on structured compressive sensing. In Proceedings of the IEEE International Conference on Communications (ICC), Sydney, Australia, 10–14 June 2014; pp. 1–6.

18. Gui, G.; Peng, W.; Adachi, F. Sub-nyquist rate ADC sampling-based compressive channel estimation. *Wirel. Commun. Mob. Comput.* **2015**, *15*, 639–648. [[CrossRef](#)]
19. Candes, E.J.; Wakin, M.B.; Boyd, S.P. Enhancing sparsity by reweighted L1 minimization. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905. [[CrossRef](#)]
20. Gu, Y.; Jin, J.; Mei, S. L0-norm constraint LMS algorithm for sparse system identification. *IEEE Signal Process. Lett.* **2009**, *16*, 774–777.
21. Gui, G.; Xu, L.; Ma, W.; Chen, B. Robust adaptive sparse channel estimation in the presence of impulsive noises. In Proceedings of the IEEE International Conference on Digital Signal Processing (DSP), Singapore, 21–24 July 2015; pp. 1–5.
22. Donoho, D.L.; Huo, X. Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inf. Theory* **2001**, *47*, 2845–2862. [[CrossRef](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).