

Article

# Norm Penalized Joint-Optimization NLMS Algorithms for Broadband Sparse Adaptive Channel Estimation

Yanyan Wang<sup>1</sup> and Yingsong Li<sup>1,2,\*</sup> 

<sup>1</sup> College of Information and Communications Engineering, Harbin Engineering University, Harbin 150001, China; wangyanyan@hrbeu.edu.cn

<sup>2</sup> National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

\* Correspondence: liyingsong@ieee.org; Tel.: +86-451-8251-9808

Received: 16 June 2017; Accepted: 23 July 2017; Published: 27 July 2017

**Abstract:** A joint-optimization method is proposed for enhancing the behavior of the  $l_1$ -norm- and sum-log norm-penalized NLMS algorithms to meet the requirements of sparse adaptive channel estimations. The improved channel estimation algorithms are realized by using a state stable model to implement a joint-optimization problem to give a proper trade-off between the convergence and the channel estimation behavior. The joint-optimization problem is to optimize the step size and regularization parameters for minimizing the estimation bias of the channel. Numerical results achieved from a broadband sparse channel estimation are given to indicate the good behavior of the developed joint-optimized NLMS algorithms by comparison with the previously proposed  $l_1$ -norm- and sum-log norm-penalized NLMS and least mean square (LMS) algorithms.

**Keywords:** NLMS algorithm; least mean square (LMS); zero attracting; joint-optimization; sparse channel estimations

## 1. Introduction

Recently, adaptive channel estimation has been extensively studied all over the world [1,2]. From these adaptive filtering frameworks, the LMS adaptive filtering algorithm has been deeply discussed for adaptive control, system identification, channel estimation applications owing to its low computational complexity and easy practical realization [3–5]. Although the LMS algorithm can effectively estimate the broadband multi-path channel, it has a sensitivity to the input signal scaling, and it is difficult to select an agreeable learning rate to achieve a stable and robust channel estimation behavior [6–11]. Subsequently, the NLMS has been proposed by using the normalization of the input training signal power in order to overcome the above addressed problem [12,13]. However, the conventional NLMS algorithm cannot perform well for dealing with sparse channel estimation.

Additionally, the broadband multi-path channel or underwater communication channel might be a sparse channel, which has been studied in recent decades [13–17]. From the measurement of the wireless channel, the channel impulse response can be regarded as sparse channel. This is to say that only a few channel impulse responses in most of the multi-path channels are dominant, while the major channel responses are zeros or their magnitudes are near-zeros. The traditional LMS- and NLMS-based channel estimation methods cannot make use of the inherent sparse properties of these broadband multi-path channels [6–13]. To utilize the sparse structures of the broadband sparse channel, compressed sensing (CS) methods have been introduced for developing various channel estimation algorithms used in sparse cases [18–20]. Although some of these CS-based sparse channel estimations can achieve robust estimation performance, these CS algorithms may have high complexity for dealing with time-varying channels or they have difficulty in constructing desired measurement

matrices with the restricted isometry property limitation [21]. Thus, simple sparse channel estimation developments have attracted much more attentions in the recent decades.

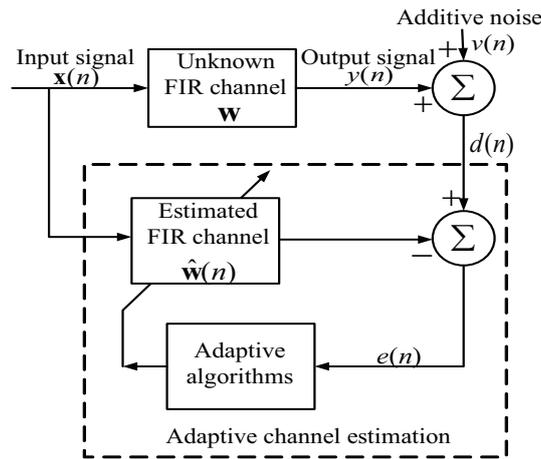
Sparse LMSs have been presented under the inspiration of the CS techniques [22,23], which are realized by using a norm constraint term in the cost function of the LMS. The first sparse LMS algorithm motivated by the CS technique is carried out by introducing a  $l_1$ -norm constraint term into the basic LMS to exploit the in-nature sparse characteristics of broadband multi-path channel [24–27]. As a result, a zero attractor is given in the updating equation of the sparse LMS algorithm to put forward a zero-attracting (ZA) LMS (ZA-LMS) algorithm. Furthermore, a reweighting ZA-LMS (RZA-LMS) was reported by using a sum-log constraint instead of the  $l_1$ -norm penalty in the ZA-LMS algorithm [24,27]. Subsequently, the zero attracting techniques have been widely researched, and a great quantity of sparse LMS algorithms was exploited by using different norm constraints, such as  $l_p$ -norm and smooth approximation  $l_0$ -norm constraints [28–32]. Furthermore, the zero attracting (ZA) technique has also been expanded into the affine projection algorithm and the normalized NLMS algorithms to further exploit the applications of the ZA algorithms [12,33–40], which includes ZA-NLMS and RZA-NLMS algorithms. However, the affine projection algorithm has higher complexity than the NLMS algorithm, which limited its applications. Thus, the sparse NLMS algorithms have been extensively studied and have been used for sparse channel estimations. However, the behavior of these NLMS-based channel estimations was affected by the modified step size and the regularized parameters. The normalized step size has an important effect on the compromise between the channel estimation behavior and convergence speed, while the regularization parameter depends on the SNR of the systems [41]. From [41], we can see that the proposed joint-optimization method cannot utilize the sparsity of the multi-channels. In addition, the step size and the regularization parameters should be selected to address the conflict requirement between the channel estimation behavior and the convergence speed.

In this paper, a joint-optimization method is discussed to enhance the behavior of the ZA- and RZA-NLMS algorithms for estimating finite impulse response (FIR) sparse channels. Their estimation behaviors are discussed via a broadband multi-path channel, which is a prior known sparse channel from the previous studies [6,7,9,12,14,24,27–40]. The proposed algorithms are realized by implementing a joint-optimization on the regularization parameter and step size to achieve a minimization of the channel estimation misalignment, where the joint-optimization method is obtained from [41]. The joint-optimization scheme is derived in detail in the context of zero attraction theory, and it is incorporated into the previously-proposed ZA- and RZA-NLMS algorithms to form zero-attracting joint-optimization NLMS (ZAJO-NLMS) and reweighted ZAJO-NLMS (RZAJO-NLMS) algorithms. The parameter selection and the behavior of our ZAJO-NLMS and RZAJO-NLMS are evaluated through a sparse channel, which is similar to [6,7,9,12,14,24,27–40]. The computer simulations indicate that our ZAJO-NLMS and RZAJO-NLMS algorithms perform better than the previously-presented ZA-NLMS, RZA-NLMS, ZA-LMS, RZA-LMS, traditional LMS and NLMS algorithms for dealing with sparse channel estimation.

The structure is illustrated herein. We review the basic NLMS and the previously-reported ZA-NLMS, as well as the RZA-NLMS algorithms in Section 2 through estimating a broadband multipath wireless communication channel. Section 3 gives the derivation of the joint-optimization scheme and the proposed ZAJO-NLMS and RZAJO-NLMS algorithms. In Section 4, our ZAJO-NLMS and RZAJO-NLMS algorithms will be evaluated through a broadband sparse multipath channel, and their channel estimation behaviors are discussed and compared with the previous ZA-NLMS, RZA-NLMS, ZA-LMS, RZA-LMS, traditional LMS and NLMS algorithms. At last, a short summary is given in Section 5.

## 2. NLMS-Based Sparse Adaptive Channel Estimation Algorithm

We review the traditional NLMS and its sparse forms, which include the ZA-NLMS and RZA-NLMS, over a sparse multipath wireless channel, which is given in Figure 1.



**Figure 1.** Typical adaptive filter-based channel estimation system. FIR, finite impulse response.

$\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-N+1)]^T$  is used as a training input signal, which is transmitted over an FIR channel  $\mathbf{w} = [w_0, w_1, \dots, w_{N-1}]^T$ . In this paper,  $N$  denotes the length of a multipath channel, and  $(\cdot)^T$  is the transpose operation. The channel output is  $y(n)$ , which is denoted as  $y(n) = \mathbf{x}^T(n)\mathbf{w}$ . The expected signal  $d(n)$  is acquired at the receiver, and  $d(n) = y(n) + v(n)$ .  $v(n)$  denotes an additional white Gaussian noise (AWGN). The NLMS-based channel estimation algorithms aim to get the unknown sparse channel  $\mathbf{w}$  by utilizing  $\mathbf{x}(n)$  and  $y(n)$  to minimize the instant error  $e(n)$ . We define  $e(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{w}}(n)$ .  $\hat{\mathbf{w}}(n)$  denotes the estimated vector. The cost function of the traditional NLMS is depicted as:

$$\begin{aligned} & \min \|\hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n)\|^2 \\ & \text{s. t. } : d(n) - \mathbf{x}^T(n)\hat{\mathbf{w}}(n+1) = 0 \end{aligned} \quad (1)$$

By using the Lagrange multiplier method to carry out the desired minimization of (1) and introducing a controlling step size, the updating of the traditional NLMS is obtained [3–5]:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{\mu e(n)\mathbf{x}(n)}{\mathbf{x}^T(n)\mathbf{x}(n) + \delta} \quad (2)$$

$\mu$  denotes a step size of the NLMS algorithm, and  $\delta$  represents a regularization parameter with small value, which is to prevent from dividing by zero. Although the NLMS algorithm can give a good estimation of the sparse channel  $\mathbf{w}$  by the use of the update Equation (2), it cannot use the sparsity property of the practical existing multi-path channels. Recently, the ZA technique has been introduced into the original cost function of the traditional NLMS for exploiting the sparseness of the channel. Then, the ZA-NLMS algorithm is presented by using an  $l_1$ -norm penalty to modify the original cost function of the traditional NLMS. Thus, the modified cost function for the reported ZA-NLMS is written as:

$$\begin{aligned} & \min \|\hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n)\|^2 + \gamma_1 \|\hat{\mathbf{w}}(n+1)\|_1, \\ & \text{s. t. } : d(n) - \mathbf{x}^T(n)\hat{\mathbf{w}}(n+1) = 0 \end{aligned} \quad (3)$$

where  $\gamma_1$  denotes a zero attracting strength parameter used to get a balance between the estimation behavior and the sparseness of  $\|\hat{\mathbf{w}}(n+1)\|_1$ . Furthermore, the Lagrange multiplier method is utilized to find out a solution of (3). As a result, the updating equation of ZA-NLMS is described as [12]:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{\mu_{\text{ZA-NLMS}} e(n)\mathbf{x}(n)}{\mathbf{x}^T(n)\mathbf{x}(n) + \delta} - \rho_{\text{ZA-NLMS}} \text{sgn}\{\hat{\mathbf{w}}(n)\}, \quad (4)$$

where  $\mu_{\text{ZA-NLMS}}$  denotes a step size and  $\rho_{\text{ZA-NLMS}}$  denotes the zero-attracting strength controlling factor for the ZA-NLMS algorithm, respectively. The ZA-NLMS algorithm utilizes the sparse characteristic of the multi-path wireless channel. However, it exerts a uniform zero attracting on all of the channel taps. Thus, the ZA-NLMS algorithm may reduce the estimation behavior when the designated channel is less sparse. To address this problem, a sum-log penalty [42] is utilized to form the RZA-NLMS algorithm, whose cost function is written as:

$$\begin{aligned} \min & \|\hat{\mathbf{w}}(n+1) - \hat{\mathbf{w}}(n)\|^2 + \gamma_2 \sum_{i=1}^N \log(1 + \varepsilon |w_i|) \\ \text{s. t.} & : d(n) - \mathbf{x}^T(n)\hat{\mathbf{w}}(n+1) = 0 \end{aligned} \quad (5)$$

We also employ the Lagrange multiplier method to solve Equation (5). Then, the channel coefficients of the RZA-NLMS algorithm are updated by: [12]

$$\hat{w}_i(n+1) = \hat{w}_i(n) + \frac{\mu_{\text{RZA-NLMS}} e(n) x_i(n)}{x_i^T(n) x_i(n) + \delta} - \rho_{\text{RZA-NLMS}} \frac{\text{sgn}\{\hat{w}_i(n)\}}{1 + \varepsilon |\hat{w}_i(n)|}, \quad (6)$$

or its vector form:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{\mu_{\text{RZA-NLMS}} e(n) \mathbf{x}(n)}{\mathbf{x}^T(n) \mathbf{x}(n) + \delta} - \rho_{\text{RZA-NLMS}} \frac{\text{sgn}\{\hat{\mathbf{w}}(n)\}}{\mathbf{1} + \varepsilon |\hat{\mathbf{w}}(n)|}, \quad (7)$$

where  $\mu_{\text{RZA-NLMS}}$  is a step size, while  $\rho_{\text{RZA-NLMS}}$  denotes a zero-attracting strength controlling factor of the RZA-NLMS algorithm, respectively. Here,  $\varepsilon$  is a threshold that is used for controlling the reweighted factor  $\mathbf{1}/(\mathbf{1} + \varepsilon |\hat{\mathbf{w}}(n)|)$ .

From the above discussions, it is observed that the previously-proposed ZA-NLMS and RZA-NLMS exert the desired zero-attractor on each iteration. The proposed zero attractors attract the zero channel coefficients to zero quickly compared with the traditional NLMS algorithm. Thus, we can say that the sparse NLMS algorithms, namely ZA-NLMS and RZA-NLMS, utilize different sparse penalties to achieve various zero attractors. The traditional NLMS-based channel estimation algorithm is concluded as follows:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \text{Error updating term}. \quad (8)$$

Comparing to the traditional NLMS method mentioned in (8), the mentioned sparse NLMS-based channel estimation algorithms provide amazing zero attractors for both ZA-NLMS and RZA-NLMS, and hence, their updated equation is:

$$\hat{\mathbf{w}}(n+1) = \underbrace{\hat{\mathbf{w}}(n) + \text{Error updating term}}_{\text{NLMS}} + \underbrace{\text{zero attractor}}_{\text{sparse NLMS}}. \quad (9)$$

### 3. Proposed Sparse Joint-Optimization NLMS Algorithms

Though the proposed sparse NLMS algorithms effectively utilize the in-nature properties of the wireless multi-path channel for achieving a superior channel estimation behavior, their performance will be affected by step size, regularization parameter  $\delta$  and the zero-attracting strength controlling factor. After that, various techniques have been presented to enhance the behavior of the ZA-NLMS and RZA-NLMS algorithms, including variable step size methods and parameter-adjusting techniques [32]. Herein, we concentrate on constructing a joint-optimization method on the regularization parameter  $\delta$  and step size.

Next, we will introduce our proposed ZAJONLMS algorithm in detail. Here, we consider the updating equation of the ZA-NLMS algorithm:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{\mu_{\text{ZA-NLMS}} e(n) \mathbf{x}(n)}{\mathbf{x}^T(n) \mathbf{x}(n) + \delta} - \rho_{\text{ZA-NLMS}} \text{sgn}\{\hat{\mathbf{w}}(n)\}, \quad (10)$$

where  $\mu_{\text{ZA-NLMS}}$  denotes a step size,  $\delta$  represents a regularization parameter and:

$$e(n) = d(n) - \mathbf{x}^T(n) \hat{\mathbf{w}}(n) \quad (11)$$

denotes the instantaneous channel estimation error at instant  $n$ . As we know,  $E\{\mathbf{x}^T(n) \mathbf{x}(n)\} = N\delta_x^2$  [41];  $E\{\cdot\}$  represents expectation, and  $\delta_x^2$  represents the variance of  $\mathbf{x}(n)$ . For large  $N$ , we can get  $\mathbf{x}^T(n) \mathbf{x}(n) \approx N\delta_x^2$  [43]. Then, we have:

$$\hat{\mathbf{w}}(n+1) \approx \hat{\mathbf{w}}(n) + \frac{\mu_{\text{ZA-NLMS}} e(n) \mathbf{x}(n)}{N\delta_x^2 + \delta} - \rho_{\text{ZA-NLMS}} \text{sgn}\{\hat{\mathbf{w}}(n)\}. \quad (12)$$

Next, we consider  $\mathbf{w}(n)$  as a channel that can be modeled as a simplified first order Markov model [41],

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mathbf{g}(n), \quad (13)$$

where  $\mathbf{g}(n)$  represents an AWGN signal with zero mean. Furthermore, we assume that  $\mathbf{g}(n)$  is independent of the channel  $\mathbf{w}(n)$ . Therefore, we can get:

$$\mathbf{R}_g = \delta_g^2 \mathbf{I}_N, \quad (14)$$

where  $\mathbf{I}_N$  is an  $N \times N$  identity matrix. Thus, the variance of  $\delta_g^2$  gives an important uncertainty on  $\mathbf{w}(n)$ . Here, we define a posteriori bias as:

$$\Delta(n) = \mathbf{w}(n) - \hat{\mathbf{w}}(n). \quad (15)$$

Here, we can get:

$$\Delta(n+1) = \Delta(n) + \mathbf{g}(n) - \frac{\mu_{\text{ZA-NLMS}} e(n) \mathbf{x}(n)}{N\delta_x^2 + \delta} + \rho_{\text{ZA-NLMS}} \text{sgn}\{\hat{\mathbf{w}}(n)\}. \quad (16)$$

Taking the  $l_2$ -norm on (16), using expectation on its left and right sides and getting rid of the uncorrelated products based on the i.i.d. assumptions, we have:

$$\begin{aligned} E\{\|\Delta(n+1)\|_2^2\} &= E\{\|\Delta(n)\|_2^2\} + N\delta_g^2 + \rho_{\text{ZA-NLMS}}^2 E\{\text{sgn}^2\{\hat{\mathbf{w}}(n)\}\} \\ &- \frac{2\mu_{\text{ZA-NLMS}}}{N\delta_x^2 + \delta} E\{\mathbf{x}^T(n) \Delta(n) e(n)\} + 2\rho_{\text{ZA-NLMS}} E\{\Delta(n) \text{sgn}\{\hat{\mathbf{w}}(n)\}\} - \frac{2\mu_{\text{ZA-NLMS}}}{N\delta_x^2 + \delta} E\{\mathbf{x}^T(n) \mathbf{g}(n) e(n)\} \\ &+ \frac{\mu_{\text{ZA-NLMS}}^2}{(N\delta_x^2 + \delta)^2} E\{e^2(n) \mathbf{x}^T(n) \mathbf{x}(n)\} \end{aligned} \quad (17)$$

Now, we concentrate on the last five terms in the Equation (17). From the above discussion, the instantaneous error is illustrated as:

$$e(n) = \mathbf{x}^T(n) \mathbf{w}(n) + \mathbf{x}^T(n) \mathbf{g}(n) + v(n) \quad (18)$$

By taking the Equation (18) into account, we can get:

$$\begin{aligned} E\{\mathbf{x}^T(n) \Delta(n) e(n)\} &= E\{\Delta^T(n) \mathbf{x}(n) e(n)\} \\ &= E\{\Delta^T(n) \mathbf{x}(n) \mathbf{x}^T(n) \Delta(n)\} = E\{\text{tr}[\Delta(n) \Delta^T(n) \mathbf{x}(n) \mathbf{x}^T(n)]\} \end{aligned} \quad (19)$$

where  $\text{tr}[\cdot]$  represents the trace operation of a matrix. Assume that the misalignments at the instant  $n$  and  $(n + 1)$  are uncorrelated. In the stable state, a posteriori bias correlation matrix is approximated as a diagonal matrix [41], which is because the bias of each coefficient tends to be uncorrelated. Thus, we obtain:

$$E \left\{ \Delta^T(n) \mathbf{x}(n) e(n) \right\} \approx \text{tr} \left\{ E \left[ \Delta(n) \Delta^T(n) \right] E \left[ \mathbf{x}(n) \mathbf{x}^T(n) \right] \right\} = \delta_x^2 E \left\{ \|\Delta(n)\|_2^2 \right\}. \quad (20)$$

Similarly, the cross-correlation  $E \left\{ \mathbf{x}^T(n) \mathbf{g}(n) e(n) \right\}$  can also be obtained based on (18). Therefore, the correlation matrix of  $\mathbf{g}(n)$  can be regarded as a diagonal matrix. By eliminating the uncorrelated terms, we obtain:

$$\begin{aligned} E \left\{ \mathbf{x}^T(n) \mathbf{g}(n) e(n) \right\} &= E \left\{ \mathbf{g}^T(n) \mathbf{x}(n) e(n) \right\} = E \left\{ \mathbf{g}^T(n) \mathbf{x}(n) \mathbf{x}^T(n) \mathbf{g}(n) \right\} \\ &= E \left\{ \text{tr} \left[ \mathbf{g}(n) \mathbf{g}^T(n) \mathbf{x}(n) \mathbf{x}^T(n) \right] \right\} \approx \text{tr} \left\{ E \left[ \mathbf{g}(n) \mathbf{g}^T(n) \right] E \left[ \mathbf{x}(n) \mathbf{x}^T(n) \right] \right\} = N \delta_x^2 \delta_g^2. \end{aligned} \quad (21)$$

Then, the expectation term  $E \left\{ e^2(n) \mathbf{x}^T(n) \mathbf{x}(n) \right\}$  can also be calculated by taking Equation (18) into consideration. Similarly, we have:

$$\begin{aligned} E \left\{ e^2(n) \mathbf{x}^T(n) \mathbf{x}(n) \right\} &= \text{tr} \left\{ E \left[ e^2(n) \mathbf{x}(n) \mathbf{x}^T(n) \right] \right\} = \text{tr} \left\{ E \left[ v^2(n) \mathbf{x}(n) \mathbf{x}^T(n) \right] \right\} + \\ &\text{tr} \left\{ E \left\{ \left[ \Delta(n) \mathbf{x}(n) + \mathbf{g}^T(n) \mathbf{x}(n) \right]^2 \mathbf{x}(n) \mathbf{x}^T(n) \right\} \right\}. \end{aligned} \quad (22)$$

Here, we assume that the correlation matrix of  $\mathbf{x}(n)$  approximates a diagonal matrix, which has been widely used for simplifying the analysis [41,44]. Furthermore, this also motivates us to further develop the second term in Equation (22) based on the Gaussian moment factoring theory [45]. Then, we obtain:

$$\begin{aligned} E \left\{ e^2(n) \mathbf{x}^T(n) \mathbf{x}(n) \right\} &= N \delta_x^2 \delta_v^2 + \text{tr} \left\{ \delta_x^4 \left\{ N \delta_g^2 + E \left\{ \|\Delta(n)\|_2^2 \right\} \right\} \mathbf{I}_N + 2 \delta_x^4 \left\{ \delta_g^2 \mathbf{I}_N + E \left[ \Delta(n) \Delta^T(n) \right] \right\} \right\} \\ &= N \delta_x^2 \delta_v^2 + (N + 2) \delta_x^4 \left\{ E \left\{ \|\Delta(n)\|_2^2 \right\} + N \delta_g^2 \right\}. \end{aligned} \quad (23)$$

As we know, at the stable status, we get:

$$E \left\{ \text{sgn} [\hat{\mathbf{w}}(n)] \right\} = \text{sgn} [\hat{\mathbf{w}}(n)], \quad (24)$$

and  $\rho_{\text{ZA-NLMS}}$  is very small. Thus, Equation (17) can be approximated to be:

$$\begin{aligned} E \left\{ \|\Delta(n+1)\|_2^2 \right\} &\approx E \left\{ \|\Delta(n)\|_2^2 \right\} + N \delta_g^2 - \frac{2\mu_{\text{ZA-NLMS}}}{N\delta_x^2 + \delta} E \left\{ \mathbf{x}^T(n) \Delta(n) e(n) \right\} \\ &- \frac{2\mu_{\text{ZA-NLMS}}}{N\delta_x^2 + \delta} E \left\{ \mathbf{x}^T(n) \mathbf{g}(n) e(n) \right\} + \frac{\mu_{\text{ZA-NLMS}}^2}{(N\delta_x^2 + \delta)^2} E \left\{ e^2(n) \mathbf{x}^T(n) \mathbf{x}(n) \right\} \end{aligned} \quad (25)$$

From the discussions above, we substitute Equations (20), (21) and (23) into Equation (25) and denote  $m(n) = E \left\{ \|\Delta(n+1)\|_2^2 \right\}$ ; we have:

$$m(n) = f(\mu_{\text{ZA-NLMS}}, \delta, N, \delta_x^2) E \left\{ \|\Delta(n)\|_2^2 \right\} + g(\mu_{\text{ZA-NLMS}}, \delta, N, \delta_x^2, \delta_v^2, \delta_g^2), \quad (26)$$

where:

$$f(\mu_{\text{ZA-NLMS}}, \delta, N, \delta_x^2) = 1 - \frac{2\delta_x^2}{N\delta_x^2 + \delta} \mu_{\text{ZA-NLMS}} + \frac{(N+2)\delta_x^4}{(N\delta_x^2 + \delta)^2} \mu_{\text{ZA-NLMS}}^2, \quad (27)$$

and:

$$g(\mu_{\text{ZA-NLMS}}, \delta, N, \delta_x^2, \delta_v^2, \delta_g^2) = \frac{N\mu_{\text{ZA-NLMS}}^2 \delta_x^2 \left[ \delta_v^2 + (N+2)\delta_x^2 \delta_g^2 \right]}{(N\delta_x^2 + \delta)^2} - \frac{2N\mu_{\text{ZA-NLMS}} \delta_x^2 \delta_g^2}{N\delta_x^2 + \delta} + N\delta_g^2. \quad (28)$$

From the result in Equation (26), we can see that the convergence speed and the misadjustment components are separated from each other. The first term of Equation (26) plays a significant role in the convergence speed of adaptive filters, which depends on  $\mu_{\text{ZA-NLMS}}$ ,  $\delta$ , input signal power and filter length. It is worth noting that the convergence speed component does not rely on the  $\delta_v^2$  and  $\delta_g^2$  of the model [41]. Thus, the noise power  $\delta_v^2$  and the uncertainties  $\delta_g^2$  do not give any effect on the convergence. We can see that fastest convergence speed can be achieved when Equation (27) reaches its minimum. Then, we have:

$$\frac{\partial f(\mu_{\text{ZA-NLMS}}, \delta, N, \delta_x^2)}{\partial \mu_{\text{ZA-NLMS}}} = 0. \quad (29)$$

Therefore, we can get:

$$\mu_{\text{H}} = \frac{\delta + N\delta_x^2}{(N+2)\delta_x^2}, \quad (30)$$

where  $\mu_{\text{H}}$  is the fastest convergence speed controlling factor. As we know, the regularization parameter  $\delta$  is small, and the length  $N$  is large. Thus,  $\mu_{\text{H}} \approx 1$  for achieving the fastest convergence speed, which is well known in [3,44]. Furthermore, the stability condition is obtained by letting  $|f(\mu_{\text{ZA-NLMS}}, \delta, N, \delta_x^2)| < 1$ , which results in:

$$0 < \mu_{\text{stable}} < 2 \frac{\delta + N\delta_x^2}{(N+2)\delta_x^2} = 2\mu_{\text{H}} \quad (31)$$

Again, by considering  $\delta = 0$  and  $N \gg 2$ , the stability conditions of the NLMS and ZA-NLMS algorithms can be obtained, which is written as  $\mu_{\text{NLMS}/\text{ZA-NLMS}}$ .

The second term of Equation (26) gives large effects on the misalignment of the proposed algorithm, which significantly depends on the noise power  $\delta_v^2$  and the uncertainties  $\delta_g^2$  of the model. With the increment of these two parameters, the misalignment is also increased [41]. The lowest misalignment is obtained when Equation (28) reaches the minimum. Furthermore, by taking the step size into account, we can get the lowest misalignment, which can be expressed as:

$$\mu_{\text{L}} = \frac{\delta_g^2(\delta + N\delta_x^2)}{\delta_v^2 + (N+2)\delta_x^2\delta_g^2}. \quad (32)$$

As we know, the broadband channel is always time-varying, and hence,  $\delta_g^2 = 0$ . In this case,  $\mu_{\text{L}} = 0$ . This is to say that the lowest misalignment is achieved as the step size approximates zero [3]. From the discussions mentioned above, we follow the optimization criterion to mimic the channel estimation misalignment to obtain the optimized sparse RZA-NLMS and ZA-NLMS algorithms, which is based on the convergence analysis above.

An ideal adaptive filtering algorithm needs low misalignment and a rapid convergence speed rate. Unluckily, the results in Equations (30) and (32) give opposite directions. Thus, we need to optimize the step size to enhance the channel estimation behavior. Furthermore, the regularization affects the behaviors on sparse NLMS algorithms. From (27), we can see that the convergence decreases when the regularization parameter  $\delta$  increases, while the misalignment in Equation (28) always increases when regularization parameter  $\delta$  decreases. Thus, we should control the step size and  $\delta$  to mimic the effects on the performance of the channel estimation algorithms. Additionally, we also follow a minimization problem with respect to the channel estimation misalignment. According to Equation (26) and assuming these two parameters are dependent on time, we can impose [41]:

$$\frac{\partial m(n)}{\partial \mu_{\text{ZA-NLMS}}} = 0, \quad \frac{\partial m(n)}{\partial \delta} = 0. \quad (33)$$

Then, we can get the same result, which is expressed as:

$$\frac{\mu_{ZA-NLMS}}{N\delta_x^2 + \delta} = \frac{m(n-1) + N\delta_g^2}{N\delta_v^2 + (N+2)\delta_x^2 [m(n-1) + N\delta_g^2]}, \quad (34)$$

which gives a joint-optimization procedure. Then, we introduce Equation (34) into Equation (10) to obtain the updating equation of our ZAJO-NLMS algorithm. Then, we can get:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{[m(n-1) + N\delta_g^2] \mathbf{x}(n)e(n)}{N\delta_v^2 + (N+2)\delta_x^2 [m(n-1) + N\delta_g^2]} - \gamma_{ZAJO} \text{sgn}\{\hat{\mathbf{w}}(n)\}. \quad (35)$$

Obversely, we should update the  $m(n)$  in (35). Then, substituting (34) into (26) results in:

$$m(n) = \left\{ 1 - \frac{\delta_x^2 [m(n-1) + N\delta_g^2]}{N\delta_v^2 + (N+2)\delta_x^2 [m(n-1) + N\delta_g^2]} \right\} [m(n-1) + N\delta_g^2]. \quad (36)$$

When  $n \rightarrow \infty$ , we have:

$$\delta_x^2 (m(\infty) + N\delta_g^2) - N(N+2)\delta_g^2\delta_x^2 [m(\infty) + N\delta_g^2] - N^2\delta_g^2\delta_v^2 = 0, \quad (37)$$

whose solution is expressed as:

$$m(\infty) = \frac{N\delta_g^2}{2} \left[ N + (N+2) \sqrt{1 + \frac{4\delta_v^2}{(N+2)^2\delta_g^2\delta_x^2}} \right]. \quad (38)$$

Similarly to the extraction of the ZAJO-NLMS, a reweighting factor is incorporating into the ZAJO-NLMS. As a result, the updating equation is obtained for realizing the RZAJO-NLMS algorithm:

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{[m(n-1) + N\delta_g^2] \mathbf{x}(n)e(n)}{N\delta_v^2 + (N+2)\delta_x^2 [m(n-1) + N\delta_g^2]} - \gamma_{RZAJO} \frac{\text{sgn}\{\hat{\mathbf{w}}(n)\}}{1 + \varepsilon |\hat{\mathbf{w}}(n)|}. \quad (39)$$

From the proposed ZAJO- and RZAJO-NLMS algorithms, we found that the regularization parameter  $\delta$  and step size are joint-optimized, while the zero attractor keeps invariable. From the above discussions, we found that there are two additional zero attraction terms, namely  $-\gamma_{ZAJO} \text{sgn}\{\hat{\mathbf{w}}(n)\}$  and  $-\gamma_{RZAJO} \frac{\text{sgn}\{\hat{\mathbf{w}}(n)\}}{1 + \varepsilon |\hat{\mathbf{w}}(n)|}$ , in the ZAJO-NLMS and RZAJO-NLMS algorithms, which are different from [41] due to the zero attractors. In this paper, the proposed ZAJO-NLMS and RZAJO-NLMS algorithms are implemented by the combination of the zero-attraction-based NLMS and the joint-optimization method in [41]. As a result, the ZAJO-NLMS and RZAJO-NLMS algorithms are constructed to deal with the sparse channel estimation, which can give better performance due to the joint-optimization and the zero attractors  $-\gamma_{ZAJO} \text{sgn}\{\hat{\mathbf{w}}(n)\}$  and  $-\gamma_{RZAJO} \frac{\text{sgn}\{\hat{\mathbf{w}}(n)\}}{1 + \varepsilon |\hat{\mathbf{w}}(n)|}$ .

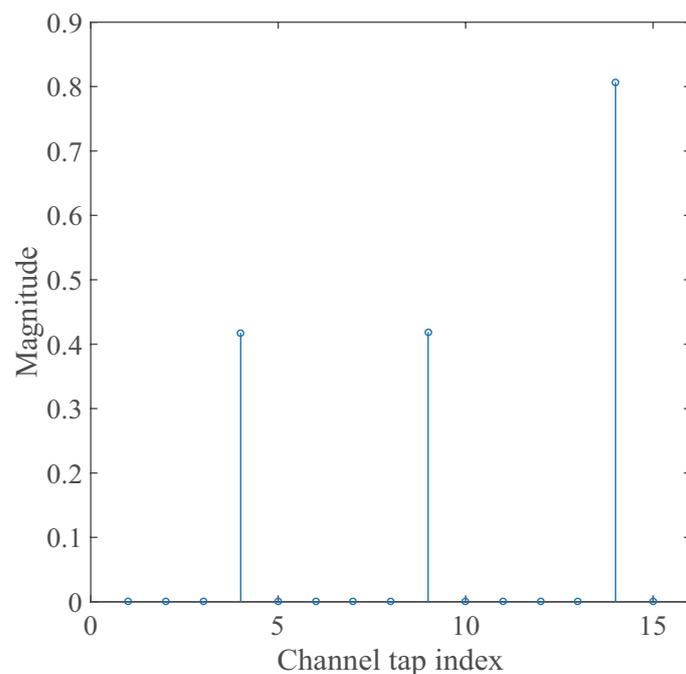
Our contributions can be summarized as follows:

- (1) Two optimized ZAJO-NLMS and RZAJO-NLMS algorithms with zero attractors have been proposed for sparse channel estimation, in the context of the state variable model.
- (2) The proposed ZAJO-NLMS and RZAJO-NLMS algorithms are realized by using the joint-optimization method and the zero attraction techniques to mimic the channel estimation misalignment.
- (3) The behaviors of the proposed ZAJO-NLMS and RZAJO-NLMS algorithms are evaluated for estimating sparse channels.

- (4) The ZAJONLMS and RZAJONLMS algorithms can achieve both faster convergence and lower misalignment than the ZA- and RZA-NLMS algorithms owing to the joint-optimization, which effectively adjusts the step size and the regularization parameter. In the future, we will develop an optimal algorithm to optimize the zero-attractor terms.

#### 4. Results and Discussion

We construct several experiments to look into the estimation behavior of our ZAJONLMS and RZAJONLMS algorithms through a multi-path wireless communication channel, which is a general sparse channel model obtained from the measurement [14,17] and which has been widely used for verifying the estimation performance of NLMS-based channel estimations [6,7,9,12,14,24,27–40]. Moreover, the channel estimation behavior is evaluated using mean-square error, and the channel estimation performance is also compared with the traditional LMS, NLMS, ZA-LMS, RZA-LMS, ZA-NLMS and RZA-NLMS algorithms. Here, a multipath channel has a length of  $N = 16$ . This channel with varying  $K$  is used for predicting the estimation behavior by means of the mean square error (MSE) standard. In the investigation, the channel estimation performance with different sparsity level  $K$  is also analyzed in detail. The  $K$  dominant coefficients are distributed randomly in the channel, and it is limited by  $\|\mathbf{w}\|_2^2 = 1$ . An example of a typical broadband sparse multi-path wireless channel is given in Figure 2.



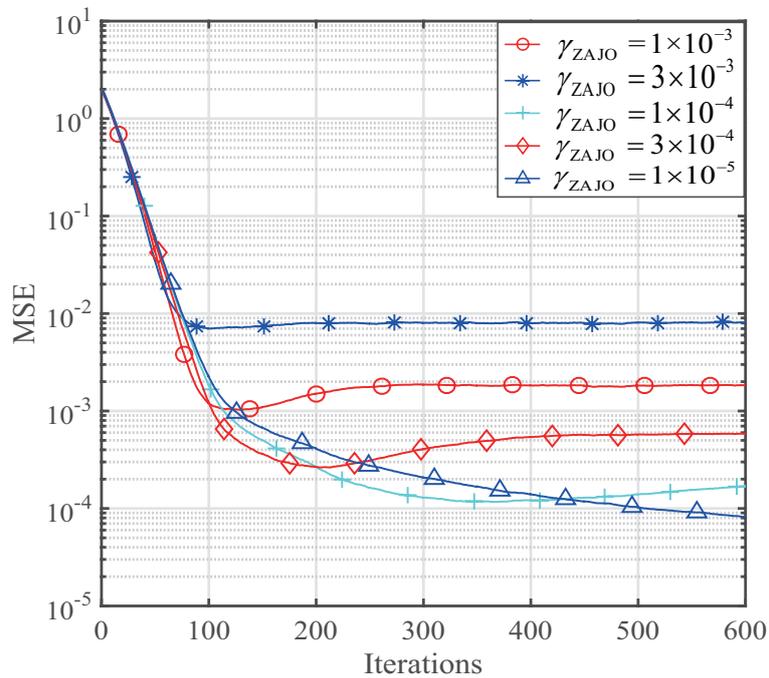
**Figure 2.** A typical broadband sparse multi-path channel.

Here, the sparse channel has a length of 16, and it has three dominant coefficients. The number of non-zero coefficients is denoted as sparsity level  $K$ . In all of the experiments,  $\mathbf{x}(n)$  is a Gaussian signal, and  $\mathbf{v}(n)$  is a Gaussian noise. The received signal has a power of  $E_b = 1$ . The power of the noise  $\mathbf{v}(n)$  is given by  $\delta_v^2$ . The estimation behaviors of our ZAJONLMS and RZAJONLMS algorithms are accessed by the MSE given by:

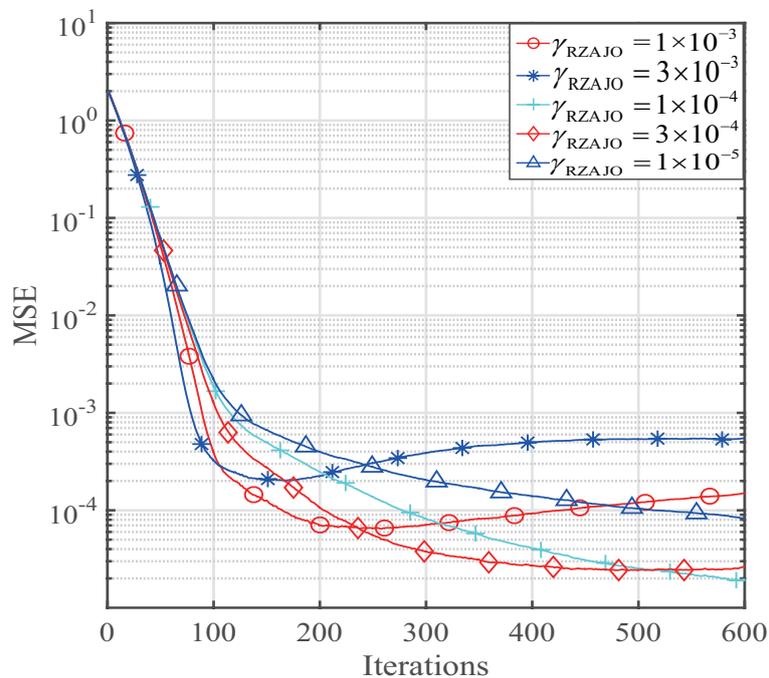
$$\text{MSE}(n) = E \left\{ \|\mathbf{w}(n) - \hat{\mathbf{w}}(n)\|^2 \right\}. \quad (40)$$

Since the step size and the regularization parameter have been optimized based on the derivation of our ZAJONLMS and RZAJONLMS, only the zero-attracting parameter can affect the performance

of these proposed algorithms. Thus, we investigated the effects of the  $\gamma_{ZAJO}$  and  $\gamma_{RZAJO}$  on the MSE. The performance is shown in Figures 3 and 4 for  $\gamma_{ZAJO}$  and  $\gamma_{RZAJO}$ , respectively.



**Figure 3.** The effect of the  $\gamma_{ZAJO}$  on the zero-attracting joint-optimization (ZAJO)-NLMS for sparse channel estimation.



**Figure 4.** Effect of the  $\gamma_{RZAJO}$  on the reweighted zero-attracting joint-optimization (RZAJO)-NLMS for sparse channel estimation.

We can see that both  $\gamma_{ZAJO}$  and  $\gamma_{RZAJO}$  have important effects on the sparse channel estimation. With a decrement of the  $\gamma_{ZAJO}$ , the MSE of the ZAJO-NLMS algorithm shown in Figure 3 is getting better when  $\gamma_{ZAJO}$  ranges from  $3 \times 10^{-3}$  to  $1 \times 10^{-4}$ . As for  $\gamma_{ZAJO} = 1 \times 10^{-5}$ , both the convergence

speed and MSE are deteriorated. From Figure 4, it is observed that  $\gamma_{RZAJ0}$  has the same trend as that of the  $\gamma_{ZAJ0}$ . Thus, we can properly select  $\gamma_{ZAJ0}$  and  $\gamma_{RZAJ0}$  to achieve desired channel estimation performance. According to the parameter effects in Figures 3 and 4, we choose  $\gamma_{ZAJ0} = \gamma_{RZAJ0} = 3 \times 10^{-4}$  to investigate the effects of the sparsity level on the sparse channel estimation performance.

Next, we set the sparsity level of the broadband sparse multi-path to  $K = 1, 2, 4, 8$  to analyze the channel estimation performance. The simulation parameters are  $\mu_{LMS} = \mu_{ZALMS} = \mu_{RZALMS} = 0.05$ ,  $\mu_{NLMS} = \mu_{ZANLMS} = \mu_{RZANLMS} = 0.5$ ,  $\rho_{ZALMS} = \rho_{RZALMS} = \rho_{ZANLMS} = \rho_{RZANLMS} = \gamma_{ZAJ0} = \gamma_{RZAJ0} = 3 \times 10^{-4}$ . The results are comparatively shown in Figures 5–8 for  $K = 1, K = 2, K = 4$  and  $K = 8$ , respectively. It is observed from Figure 5 that our RZAJ0-NLMS algorithm achieves the quickest convergence and smallest steady-state error compared with the traditional LMS, NLMS, ZA-LMS, RZA-LMS, ZA-, RZA- and ZAJ0 NLMS algorithms. The proposed RZAJ0-NLMS algorithm has much more gains than the traditional RZA-NLMS algorithm. Additionally, our proposed ZAJ0-NLMS algorithm achieves better channel estimation behaviors than those of the RZA-NLMS algorithm for  $K = 1$ . When the sparsity level  $K$  increases from 2 to 8, the gain between the RZAJ0-NLMS and RZA-NLMS algorithm is getting small. However, our proposed RZAJ0-NLMS algorithm outperforms all of the other sparse channel estimation algorithms with reference to both the MSE and convergence. When  $K = 8$ , our proposed ZAJ0-NLMS algorithm achieves quicker convergence compared to the traditional NLMS to get the same MSE level. On the contrary, the behavior of the ZA-NLMS becomes worse than that of the traditional NLMS. Thereby, our proposed RZAJ0- and ZAJ0-NLMS algorithms are more useful for adaptive sparse channel estimation applications.

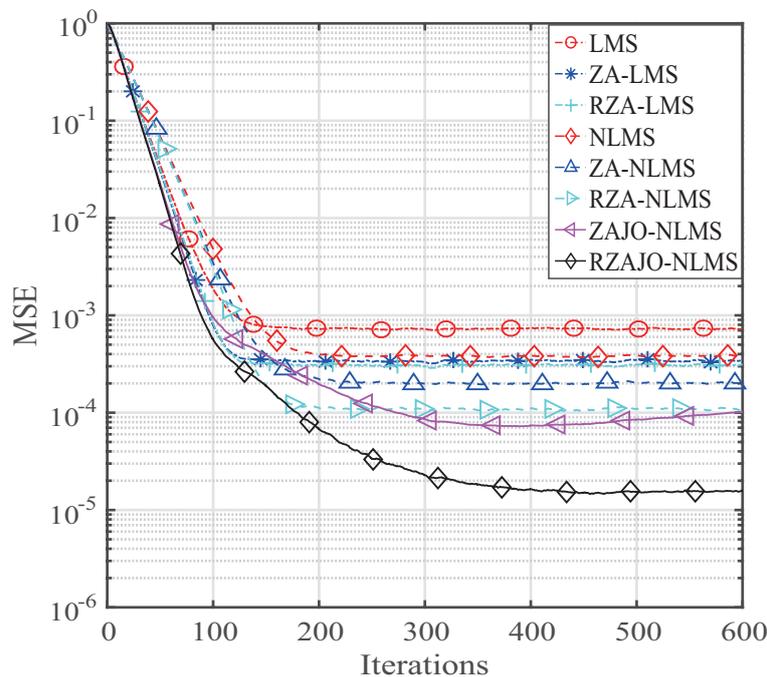


Figure 5. MSE of our presented ZAJ0- and RZAJ0-NLMS algorithms for  $K = 1$ .

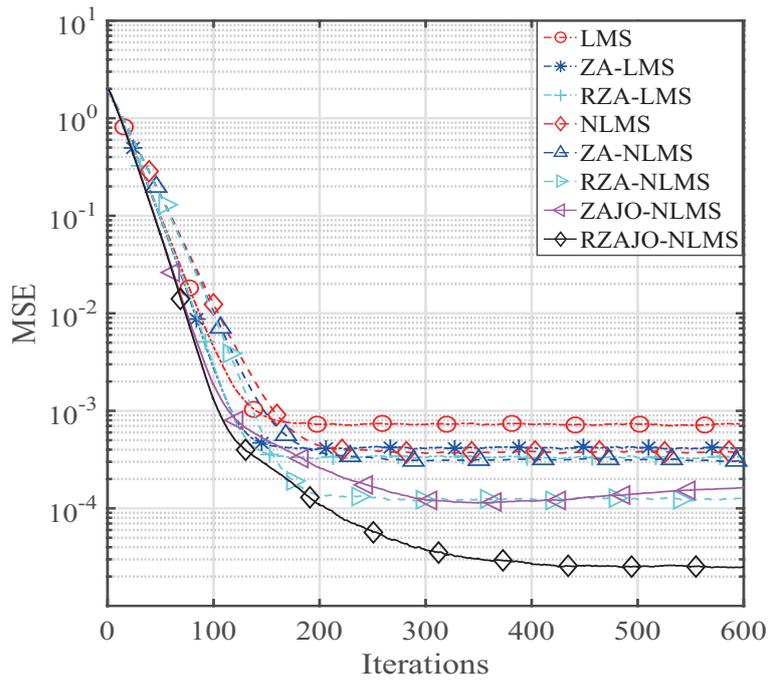


Figure 6. MSE of our presented ZAJO- and RZAJO-NLMS algorithms for  $K = 2$ .

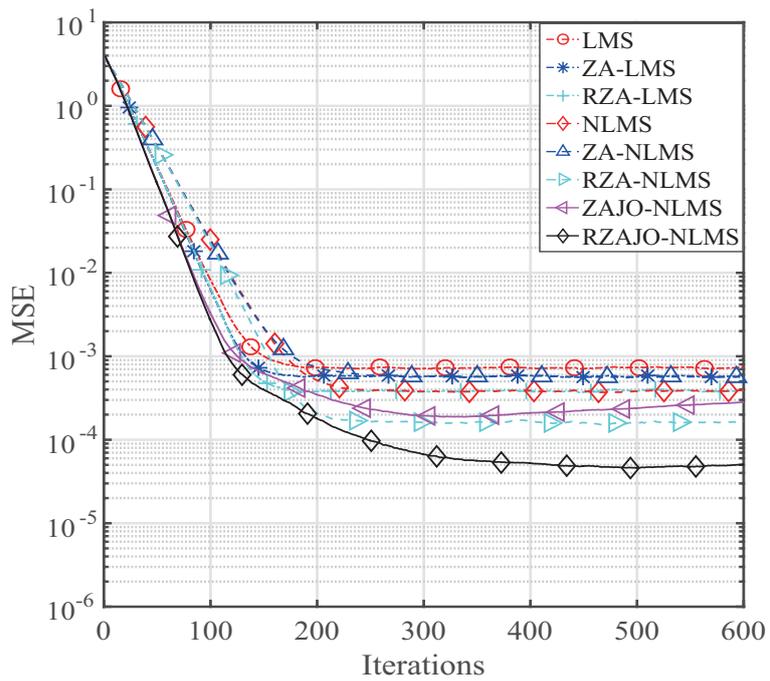


Figure 7. MSE of our presented ZAJO- and RZAJO-NLMS algorithms for  $K = 4$ .

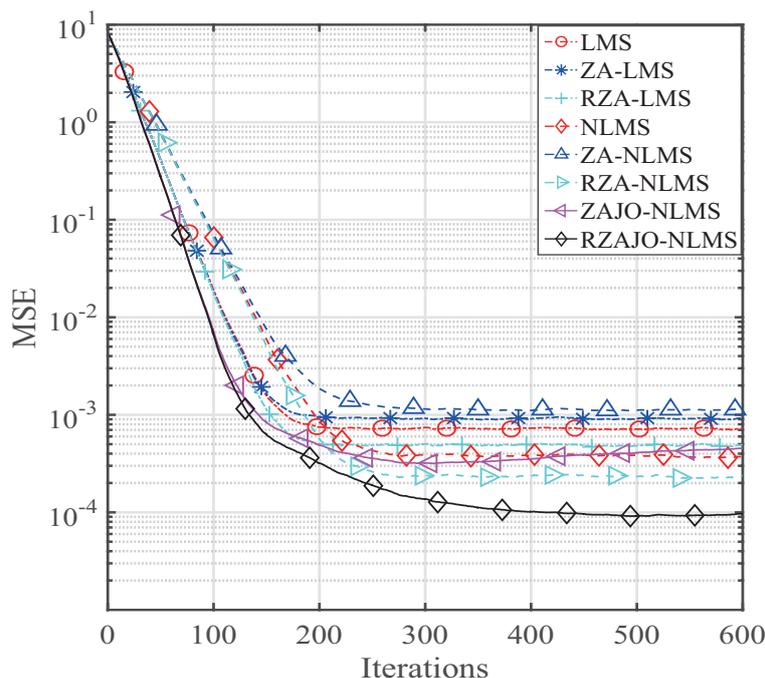


Figure 8. MSE of our presented ZAJO- and RZAJONLMS algorithms for  $K = 8$ .

## 5. Conclusions

We proposed joint-optimization sparse NLMS algorithms, namely RZAJONLMS and ZAJO-NLMS. The joint-optimization was realized by using a state model to improve the channel estimation performance of both the ZA- and RZA-NLMS algorithms. Our RZAJONLMS and ZAJO-NLMS algorithms are based on the joint-optimization of step size and the regularization parameter. The proposed joint-optimization was derived in detail. Furthermore, the estimation behavior of our RZAJONLMS and ZAJO-NLMS algorithms is evaluated on a broadband sparse multi-path channel with different sparsity levels. The results verified that our RZAJONLMS algorithm provides the fastest convergence speed rate and lowest MSE. In addition, the proposed ZAJO-NLMS outperforms the previously-reported ZA-NLMS and traditional NLMS algorithms.

This study provided the RZAJONLMS and ZAJO-NLMS algorithms based on the zero attracting and the joint-optimization techniques. The proposed joint-optimization technique can be expanded to the  $l_p$ -norm ( $0 \leq p \leq 1$ ) constrained NLMS, normalized LMF (NLMF) and normalized least mean mixed norm (NLMMN) to enhance the sparse channel estimation performance or sparse system identification. In addition, the proposed method can be used for exploiting the two-dimensional (2D) adaptive filters for imaging processing, which can be used for medical imaging denoising applications [46]. Moreover, our proposed RZAJONLMS and ZAJO-NLMS algorithms can be integrated into the orthogonal frequency-division multiplexing (OFDM) and multiple-input multiple-output (MIMO) OFDM systems to improve the quality of the communication systems [47,48].

**Acknowledgments:** This work was partially supported by the National Key Research and Development Program of China-Government Corporation Special Program (2016YFE0111100), the National Science Foundation of China (61571149), the Science and Technology innovative Talents Foundation of Harbin (2016RAXXJ044), the Projects for the Selected Returned Overseas Chinese Scholars of Heilongjiang Province and Ministry of Human Resources and Social Security of the People's Republic of China (MOHRSS) and the PhD Student Research and Innovation Fund of the Fundamental Research Funds for the Central Universities (HEUGIP201707).

**Author Contributions:** Yanyan Wang wrote the draft of the paper, wrote the code and did the simulations of this paper. Yingsong Li helped to check the coding and simulations, and he also put forward the idea of the RZAJONLMS and ZAJO-NLMS algorithms. He also provided the analysis of the paper. Both of the authors wrote this paper together, and they have read and approved the final manuscript.

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

## References

1. Korowajczuk, L. *LTE, WiMAX and WLAN Network Design, Optimization and Performance Analysis*; John Wiley: Chichester, West Susses, UK, 2011.
2. Proakis, J.G.; Salehi, M. *Digital Communications*, 5th ed.; McGraw-Hill: New York, NY, USA, 2001.
3. Haykin, S.; Widrow, B. *Least-Mean-Square Adaptive Filters*; Wiley: Hoboken, NJ, USA, 2003.
4. Diniz, P.S.R. *Adaptive Filtering Algorithms and Practical Implementation*, 4th ed.; Springer: New York, NY, USA, 2013.
5. Haykin, S. *Adaptive Filter Theory*, 4th ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2002.
6. Homer, J.; Mareets, I.; Hoang, C. Enhanced detection-guided NLMS estimation of sparse FIR-modeled signal channels. *IEEE Trans. Circuits Syst. I* **2006**, *53*, 1783–1791.
7. Morelli, M.; Sanguinetti, L.; Mengali, U. Channel estimation for adaptive frequency-domain equalization. *IEEE Trans. Wirel. Commun.* **2005**, *4*, 2508–2518.
8. Vega, L.R.; Rey, H.G.; Benesty, J.; Tressens, S. A stochastic model for a new robust NLMS algorithm. In Proceedings of the 2007 15th European Signal Processing Conference (EUSIPCO'07), Poznan, Poland, 3–7 September 2007.
9. Wang, Y.; Li, Y.; Yang, R. Sparsity-aware adaptive filter algorithms with mixed controlled  $l_2$  and  $l_p$  error criterions for sparse adaptive channel estimation. Available online: <https://doi.org/10.1016/j.jfranklin.2017.07.036> (accessed on 24 July 2017).
10. Li, Y.; Wang, Y.; Jiang, T. Sparse least mean mixed-norm adaptive filtering algorithms for sparse channel estimation applications. *Int. J. Commun. Syst.* **2017**, *30*, 1–14.
11. Wang, Y.; Li, Y. Sparse multi-path channel estimation using norm combination constrained set-membership NLMS algorithms. *Wirel. Commun. Mob. Comput.* **2017**, *2017*, 1–10.
12. 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'13), Shanghai, China, 7–10 April 2013.
13. Li, Y.; Wang, Y.; Yang, R.; Albu, F. A soft parameter function penalized normalized maximum correntropy criterion algorithm for sparse system identification. *Entropy* **2017**, *19*, 45.
14. Cotter, S.F.; Rao, B.D. Sparse channel estimation via matching pursuit with application to equalization. *IEEE Trans. Commun.* **2002**, *50*, 374–377.
15. Vuokko, L.; Kolmonen, V.M.; Salo, J.; Vainikainen, P. Measurement of large-scale cluster power characteristics for geometric channel models. *IEEE Trans. Antennas Propag.* **2007**, *55*, 3361–3365.
16. Gendron, P.J. Shallow water acoustic response and platform motion modeling via a hierarchical Gaussian mixture model. *J. Acoust. Soc. Am.* **2016**, *139*, 1923–1937.
17. Maechler, P.; Greisen, P.; Sporrer, B.; Steiner, S.; Felber, N.; Brug, A. Implementation of greedy algorithms for LTE sparse channel estimation. In Proceedings of the 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers (ASILOMAR'10), Pacific Grove, CA, USA, 7–10 November 2010; pp. 400–405.
18. Meng, J.; Yin, W.; Li, Y.; Nguyen, N.T.; Zhu, H. Compressive sensing based high-resolution channel estimation for OFDM system. *IEEE J. Sel. Top. Signal Process.* **2012**, *6*, 15–25.
19. Bajwa, W.B.; Haupt, J.; Sayeed, A.M.; Nowak, R. Compressed channel sensing: A new approach to estimating sparse multipath channels. *Proc. IEEE* **2010**, *98*, 1058–1076.
20. Taub{ock, G.; Hlawatach, F. A compressed sensing technique for OFDM channel estimation in mobile environment: exploiting channel sparsity for reducing pilots. In Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'08), Las Vegas, NV, USA, 31 March–4 April 2008; pp. 2885–2888.
21. Candés, E.J. The restricted isometry property and its implications for compressed sensing. *C. R. Math.* **2008**, *346*, 589–592.
22. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **1996**, *58*, 267–288.
23. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306.

24. Chen, Y.; Gu, Y.; Hero, A.O. Sparse LMS for system identification. In Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'09), Taipei, Taiwan, 19–24 April 2009; pp. 3125–3128.
25. Li, Y.; Jin, Z.; Wang, Y. Adaptive channel estimation based on an improved norm constrained set-membership normalized least mean square algorithm. *Wirel. Commun. Mob. Comput.* **2017**, *2017*, 1–8.
26. Li, Y.; Jin, Z.; Wang, Y.; Yang, R. A robust sparse adaptive filtering algorithm with a correntropy induced metric constraint for broadband multi-path channel estimation. *Entropy* **2016**, *18*, 380.
27. Shi, K.; Shi, P. Convergence analysis of sparse LMS algorithms with  $l_1$ -norm penalty based on white input signal. *Signal Process.* **2010**, *90*, 3289–3293.
28. Wu, F.Y.; Tong, F. Gradient optimization  $p$ -norm-like constraint LMS algorithm for sparse system estimation. *Signal Process.* **2013**, *93*, 967–971.
29. Taheri, O.; Vorobyov, S.A. Sparse channel estimation with  $L_p$ -norm and reweighted  $L_1$ -norm penalized least mean squares. In Proceedings of the 36th International Conference on Acoustics, Speech and Signal Processing (ICASSP'11), Prague, Czech Republic, 22–27 May 2011; pp. 2864–2867.
30. Gu, Y.; Jin, J.; Mei, S.  $L_0$  norm constraint LMS algorithms for sparse system identification. *IEEE Signal Process. Lett.* **2009**, *16*, 774–777.
31. Niazadeh, R.; Ghalehjeh, S.H.; Babaie-Zadeh, M.; Jutten, C. ISI sparse channel estimation based on SL0 and its application in ML sequence-by-sequence equalization. *Signal Process.* **2012**, *92*, 1875–1885.
32. Li, Y.; Hamamura, M. Zero-attracting variable-step size least mean square algorithms for adaptive sparse channel estimation. *Int. J. Adap. Control Signal Process.* **2015**, *29*, 1189–1206.
33. Meng, R.; Lamare, R.C.; Nascimento, V.H. Sparsity-aware affine projection adaptive algorithms for system identification. In Proceedings of the Sensor Signal Processing for Defence (SSPD 2011), London, UK, 27–29 September 2011; pp.1–5.
34. Lima, M.V.S.; Martins, W.A.; Diniz, P.S.Z. Affine projection algorithms for sparse system identification. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'13), Vancouver, BC, Canada, 26–31 May 2013; pp. 5666–5670.
35. Li, Y.; Hamamura, M. An improved proportionate normalized least-mean-square algorithm for broadband multipath channel estimation. *Sci. World J.* **2014**, *2014*, 1–9.
36. Li, Y.; Hamamura, M. Smooth approximation  $l_0$ -norm constrained affine projection algorithm and its applications in sparse channel estimation. *Sci. World J.* **2014**, *2014*, 1–14.
37. Li, Y.; Li, W.; Yu, W.; Wan, J.; Li, Z.; Sparse adaptive channel estimation based on  $l_p$ -norm-penalized affine projection algorithm. *Int. J. Antennas Propag.* **2014**, *2014*, 1–8.
38. Li, Y.; Zhang, C.; Wang, S. Low-complexity non-uniform penalized affine projection algorithm for sparse system identification. *Circuits Syst. Signal Process.* **2016**, *35*, 1611–1624.
39. Li, Y.; Wang, Y.; Jiang, T. Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo cancelation. *AEU-Int. J. Electron. Commun.* **2016**, *70*, 895–902.
40. Li, Y.; Wang, Y.; Jiang, T. Norm-adaption penalized least mean square/fourth algorithm for sparse channel estimation. *Signal Process.* **2016**, *128*, 243–251.
41. Ciochina, S.; Paleologu, C.; Benesty, J. An optimized NLMS algorithm for system identification. *Signal Process.* **2016**, *118*, 115–121.
42. Candés, E.J.; Wakin, M.B.; Noyd, S.P. Enhancing sparsity by reweighted  $l_1$ -minimization. *J. Fourier Anal. Appl.* **2008**, *15*, 877–905.
43. Benesty, J.; Rey, H.; ReyVega, L.; Tressens, S. A nonparametric VSS-NLMS algorithm. *IEEE Signal Process. Lett.* **2006**, *13*, 581–584.
44. Isserlis, L. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika* **1918**, *12*, 134–139.
45. Sulyman, A.I.; Zerguine, A. Convergence and steady-state analysis of a variable step size NLMS algorithm. *Signal Process.* **2003**, *83*, 1255–1273.
46. Rodger, J.A. Discovery of medical big data analytics: improving the prediction of traumatic brain injury survival rates by datamining patient informatics processing software hybrid hadoop hive. *Inform. Med. Unlocked* **2015**, *1*, 17–26.

47. Elangovan, K. Comparative study on the channel estimation for OFDM system using LMS, NLMS and RLS algorithms. In Proceedings of the 2012 International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME), Tamilnadu, India, 21–23 March 2012.
48. Gui, G.; Adachi, F. Adaptive sparse channel estimation for time-variant MIMO-OFDM systems. In Proceedings of the 2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC), Sardinia, Italy, 1–5 July 2013.



© 2017 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/>).