Research Article

Sparse Multipath Channel Estimation Using Norm Combination Constrained Set-Membership NLMS Algorithms

Yanyan Wang and Yingsong Li

College of Information and Communications Engineering, Harbin Engineering University, Harbin 150001, China

Correspondence should be addressed to Yingsong Li; liyingsong@ieee.org

Received 22 December 2016; Accepted 16 February 2017; Published 28 February 2017

Copyright © 2017 Yanyan Wang and Yingsong Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A norm combination penalized set-membership NLMS algorithm with $l_0$ and $l_1$ independently constrained, which is denoted as $l_0$ and $l_1$ independently constrained set-membership (SM) NLMS (L0L1SM-NLMS) algorithm, is presented for sparse adaptive multipath channel estimations. The L0L1SM-NLMS algorithm with fast convergence and smallest estimation error is implemented by independently exerting penalties on the channel coefficients via controlling the large group and small group channel coefficients which are implemented by $l_0$ and $l_1$ norm constraints, respectively. Additionally, a further improved L0L1SM-NLMS algorithm denoted as reweighted L0L1SM-NLMS (RL0L1SM-NLMS) algorithm is presented via integrating a reweighting factor into our L0L1SM-NLMS algorithm to properly adjust the zero-attracting capabilities. Our developed RL0L1SM-NLMS algorithm provides a better estimation behavior than the presented L0L1SM-NLMS algorithm for implementing an estimation on sparse channels. The estimation performance of the L0L1SM-NLMS and RL0L1SM-NLMS algorithms is obtained for estimating sparse channels. The achieved simulation results show that our L0L1SM- and RL0L1SM-NLMS algorithms are superior to the traditional LMS, NLMS, SM-NLMS, ZA-LMS, RZA-LMS, and ZA-, RZA-, ZASM-, and RZASM-NLMS algorithms in terms of the convergence speed and steady-state performance.

1. Introduction

Broadband communication technology has attracted much attention for modern wireless communications, and signal transmission on such a broadband wireless channel usually encounters a frequency selecting phenomenon [1–5]. Moreover, the frequency-selective behavior often leads to sparse properties in such broad channels [1, 6]. In those broadband wireless multipath communication channels, most of these unknown channels are often sparse [1–3], in which only a fewer coefficients are large in magnitude. In the sparse channel, only fewer channel taps are dominant, indicating that a number of channel taps are close to zeros or zeros. On the other hand, signal though such frequency selecting sparse channels and noise environments may reduce the communication quality. Therefore, precision channel estimation methods are desired to recovery unknown channel which is often illustrated by finite impulse response (FIR) [1–3]. Adaptive filter techniques are regarded to be useful channel estimation methods owing to their fast convergence speed and good estimation behavior [7–9]. After that, various adaptive channel estimation algorithms were presented in order to guarantee the stable propagation and effective signal transmission [8–13], such as LMS and SM-NLMS algorithms [10–13]. Although these adaptive filtering algorithms can achieve robust estimation performance, they cannot deal with the sparse channel estimation well.

In order to make use of the prior sparseness characteristics of the multipath broadband channels, several sparse LMS algorithms were reported by modifying the cost function of the conventional LMS algorithms [14–18], which is inspired by the compressive sensing (CS) [19, 20]. In [14], a sparse LMS algorithm has been reported by using the modified cost function via a $l_1$-norm penalty, resulting in zero-attraction (ZA) LMS (ZA-LMS) algorithm. ZA-LMS can provide a good behavior for sparse adaptive channel estimations by exerting a zero-attraction on all channel taps. The estimation behavior of the ZA-LMS may be reduced...
because of the uniform penalty. Then, a reweighted ZA-LMS (RZA-LMS) algorithm was presented by considering a log-sum function as a constraint of the channel coefficient vector [14]. Subsequently, the RZA-LMS achieved an improved channel estimation behavior in accordance with convergence and estimation misalignment. After that, the zero-attracting technique has been realized by using $l_0$-norm [21], smooth approximation $l_p$-norm [15], and nonuniform constraints [18] to further develop the sparse properties in broadband channel. However, their estimation behavior may be degraded since the LMS is sensitive to the input signal scaling. Then, the zero-attracting methods were expanded to the NLMS [22], LMF [23], LMS/F [24–27], leaky LMS [28], and affine projection algorithms [29–33]. However, some of them are high in complexity and others need an extra balance parameter to adjust the combination of the LMF and LMS algorithms.

Recently, the zero-attracting methods and set-membership filtering theory have been used to take use of the sparseness of the sparse channel to reduce computational complexity [10, 34, 35]. In [35], the ZA and RZA methods are further investigated on the SM-NLMS to develop RZASM- and ZASM-NLMS algorithms. As a result, the RZASM- and ZASM-NLMS algorithms can provide a better channel estimation behavior than NLMS and its variants. However, these two sparse SM-NLMS adaptive filtering algorithms cannot effectively adjust the zero-attracting according to the channel coefficients in real time.

We propose a norm combination penalized SM-NLMS algorithm with $l_0$ and $l_1$ independently constrained, which is realized by employing a constraint on the channel taps according to the value of the channel coefficients and it is named as $l_0$ and $l_1$ independently constrained set-membership NLMS (L0L1SM-NLMS) algorithm. The realization of our L0L1SM-NLMS is devised via integrating $l_0$-norm and $l_1$-norm into the SM-NLMS's cost function for constructing a desired zero-atraction term which separates the channel taps into small and large groups. Then, the two different groups are attracted to zero based on $l_0$-norm and $l_1$-norm penalties. The proposed L0L1SM-NLMS algorithm is given in detail. Also, a reweighting method is introduced into the zero-atraction term to enhance the robustness of the L0L1SM-NLMS algorithm, resulting in a reweighting L0L1SM-NLMS (RL0L1SM-NLMS) algorithm. We give an evaluation of our developed L0L1SM-NLMS and RL0L1SM-NLMS algorithms in designated sparse channels. The obtained results by estimating sparse channels illustrate that our L0L1SM- and RL0L1SM-NLMS algorithms are more excellent compared with the LMS, NLMS, ZA-LMS, RZA-LMS, and ZA-, RZA-, SM-, ZASM-, and RZASM-NLMS algorithms as for the convergence and estimation misalignment.

The rest of the paper is given herein. Section 2 gives a review of the conventional SM-NLMS and SM filtering theory. Also, the previously proposed ZASM-NLMS algorithm is discussed. In Section 3, our developed L0L1SM- and RL0L1SM-NLMS algorithms are derived thoroughly. In Section 4, channel estimation behavior of our L0L1SM-NLMS and RL0L1SM-NLMS is investigated and discussed in detail. The last section is conclusion of this paper.

2. Conventional SM-NLMS Algorithm

2.1. SM Filtering Theory. A training signal $x(n) = [x(n - 1), x(n), \ldots, x(n - N)]^T$ is considered to give a discussion on the SM-NLMS. An AWGN signal $v(n)$ and an expected signal $d(n)$ are used to discuss the SM filtering (SMF) theory and typical adaptive channel estimation (ACE) system. $x(n)$ is conveyed to an unknown FIR wireless communication channel $w$ and the multipath fading channel's output is $y(n)$ that is gotten by $y(n) = x^T(n)w$. At receiver side, the expected signal $d(n)$ is contaminated by $v(n)$. The purpose of the ACE aims to minimize the estimation error $e(n)$ which denotes a difference between $y(n)$ and $d(n)$. Thus, we can get $e(n) = d(n) - x^T(n)\tilde{w}(n)$.

The typical adaptive filter (AF) algorithms are employed to get an estimation of the unknown FIR channel $w$ via minimizing an error function related to estimation error $e(n)$. For instance, the LMS algorithm uses the second-order estimation error $e^2(n)$. Moreover, the NLMS algorithm utilizes normalized power of $x(n)$ to improve the estimation behavior of the LMS. As for SMF theory, a certain bound is exerted on $e(n)$ [10–13]. The SMF methods employ an interested subspace to create a model $e(n)$. Assume that a model space $S$ is comprised of input-vector-desired-output pairs (IVDOPs). In SMF theory, an error criterion is used for bounding $S$. Parameter estimations are bounded based on a parameter $\gamma$ for all the data in $S$. Thereby, SMF algorithm should properly choose a special set in the parameter space and hence it is different with point estimation. According to the SMF principle, we have [10–13]

$$|e(n)|^2 \leq \gamma^2 \quad \forall (d(n), x(n)) \in S,$$

where $(d(n), x(n))$ denotes IVDOPs. As arbitrary $\forall (d(n), x(n)) \in S$, we have a solution of the possible vectors $w$ by the use of the following equation [10–13]:

$$\Theta = \bigcap_{(d(n), x(n)) \in S} \left\{ \tilde{w} \in \mathbb{R}^N : \left|d(n) - x^T(n)\tilde{w}(n)\right| \leq \gamma \right\},$$

where $\mathbb{R}^N$ is a vector space with a dimension of $N$. If $k$ IVDOPs $\{x_i(n), d_i(n)\}_{i=1}^k$ are used for training the filter, the measurement set of the SMF algorithms can be written as [10–13]

$$H_k = \{ \tilde{w} \in \mathbb{R}^N : \left|d_k(n) - x_k^T(n)\tilde{w}(n)\right| \leq \gamma \}.$$  

The SMF algorithm is to find the solutions following an exact set $\Psi_k$ which possesses $k$ observed IVDOPs

$$\Psi_k = \bigcap_{i=1}^k H_i.$$  

In fact, $\Theta$ is a subset of $\Psi_k$ in each iteration.
2.2. SM-NLMS Algorithm. Based on SMF theory and NLMS algorithm, the SM-NLMS has been developed by finding a minimization of $\|\hat{w}(n+1) - \tilde{w}(n)\|^2$ within a constraint of $\tilde{w}(n+1) \in H_k$ [10]. The SM-NLMS algorithm finds out the solution of the following optimization for $\tilde{w}(n) \notin H_k$

$$\min \|\hat{w}(n+1) - \tilde{w}(n)\|^2$$
$$\text{s.t.} \quad d(n) - x^T(n) \tilde{w}(n+1) = \gamma.$$  

(5)

We use Lagrange multiplier method to find out the minimization of (5). Therefore, the SM-NLMS's updating equation is

$$\tilde{w}(n+1) = \tilde{w}(n) + \frac{x(n)}{x^T(n)x(n) + \delta} \alpha e(n), \quad (6)$$

where

$$\alpha = \begin{cases} 1 - \frac{\gamma}{|e(n)|} & \text{for } |e(n)| > \gamma \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where $\delta > 0$ is to avoid dividing by zero. From (6), we can see that the SM-NLMS's updating formula is similar to basic NLMS algorithm. Herein, $\alpha$ acts as a step factor.

Inspired by the CS and ZA techniques, sparse SM-NLMS algorithm denoted as ZASM-NLMS was presented by exerting a $l_1$-norm constraint on the SM-NLMS's cost function, and hence, the ZASM-NLMS solves the following problem, given as [35]

$$\min \|\hat{w}(n+1) - \tilde{w}(n)\|^2 + \gamma_1 \|\tilde{w}(n+1)\|_1$$
$$\text{s.t.} \quad d(n) - x^T(n) \tilde{w}(n+1) = \gamma,$$  

(8)

where $\gamma_1$ denotes a ZA parameter. Similarly, we employ the Lagrange multiplier method to acquire the updating equation of the reported ZASM-NLMS whose updating equation is

$$\tilde{w}(n+1) = \tilde{w}(n) + \frac{x(n)}{x^T(n)x(n) + \delta} \alpha e(n) \quad (9)$$

where $G(n) = (x(n)x^T(n)/x^T(n)x(n) - 1)$. Since the updating equation for ZASM-NLMS is complex, a simple method can be employed to solve (8), which brings about an optimization problem without constraint. For the sake of compatibility with the traditional SM-NLMS, we can use the following equation to get the solution of (8) [35]:

$$\min \frac{1}{2} \|\hat{w}(n+1) - \tilde{w}(n)\|^2$$
$$+ \frac{1}{2} \|e(n) - \gamma \text{ZASL} \|_{R(n)}^2 + \gamma_1 \|\tilde{w}(n+1)\|_1, \quad (10)$$

where $R(n) = [x^T(n)x(n)]^{-1}$, and let $\|e(n) - \gamma \text{ZASL} \|_{R(n)}^2 = [e(n) - \gamma \text{ZASL}]^T R(n) [e(n) - \gamma \text{ZASL}] = 0$; we get the ZASM-NLMS's updating equation

$$\tilde{w}(n+1) = \tilde{w}(n) + \frac{x(n)}{x^T(n)x(n) + \delta} \alpha e(n)$$
$$- \rho \text{sgn}[\tilde{w}(n)], \quad \rho > 0$$  

(11)

where $\rho > 0$ denotes a ZA ability factor to give a balance between the estimation misalignment and the sparse constraint of $\tilde{w}(n)$ and $\text{sgn}[\cdot]$ represents a sign function with a component-wise implementation, and it is given by [14]

$$\text{sgn}[x] = \begin{cases} \frac{x}{|x|} & x \neq 0 \\ 0 & x = 0. \end{cases} \quad (12)$$

In contrast to the update equation in (6), it is found that the ZASM-NLMS provides an extra term $-\rho \text{sgn}[\tilde{w}(n)]$ in its update equation, which is to provide a fast attraction to force the small coefficients in magnitude to zero rapidly. Moreover, the ZA strength is controlled by parameter $\rho$ and the ZASM-NLMS algorithm provides the same zero-attraction on all the channel coefficients. As a result, the ZASM-NLMS algorithm cannot effectively distinguish the zero and nonzero channel coefficients, which may degrade its performance for less sparse systems. To improve its performance, an improved sparse SM-NLMS is presented by using a sum-log function to replace the $l_1$-norm in ZASM-NLMS algorithm, which was named as RZASM-NLMS algorithm. Although the RZASM-NLMS algorithm improves the performance of the ZASM-NLMS, it also increases the computational complexity.

3. Proposed L0L1SM-NLMS Algorithms

As is known to us, the previously presented ZASM- and RZASM-NLMS algorithms estimate the sparse channel by using an $l_1$-norm and sum-log function penalty to achieve good channel estimation performance. However, they may neglect the sparse structure in-nature unknown channel in practical applications. Inspired by the CS and the ZA techniques [14–20], we propose an adaption-norm penalized SM-NLMS algorithm by exerting $l_0$-norm and $l_1$-norm on the general cost function of the basic SM-NLMS to construct a desired independent norm penalty. The proposed algorithms are to produce desired zero-attraction terms which divides the filter coefficients into large and small groups. By dynamically assigning the $l_0$-norm and $l_1$-norm penalties to the different channel taps, the proposed algorithms can attract the large and small group channel taps to zero based on $l_0$-norm and $l_1$-norm penalties. Here, a mixture of $l_0$-norm and $l_1$-norm penalty is implemented by using an $l_p$-norm penalty. The $l_p$-norm is described as

$$\|\tilde{w}(n)\|_p = \sum_{i=0}^{N-1} |\tilde{w}_i(n)|^p, \quad (13)$$

where we have a definition of $0 \leq p \leq 1$. Then, $l_0$-norm and $l_1$-norm are

$$\lim_{p \rightarrow 0} \|\tilde{w}(n)\|_p^p = \|\tilde{w}(n)\|_0, \quad (14)$$

$$\lim_{p \rightarrow 1} \|\tilde{w}(n)\|_p^p = \|\tilde{w}(n)\|_1. \quad (15)$$

We note that $\|\tilde{w}(n + 1)\|_0$ can count the nonzero channel coefficients in the sparse channel. To utilize of the advantages...
of the $l_0$-norm and $l_1$-norm, the proposed adaption-norm penalty SM-NLMS algorithm is given by

$$\text{minimize} \quad \| \hat{w} (n + 1) - \hat{w} (n) \|^2 + \gamma_{\text{pro}} \| \hat{w} (n + 1) \|^p \quad \text{subject to} \quad d (n) - x^T (n) \hat{w} (n + 1) = \gamma,$$

where $\gamma_{\text{pro}}$ is a ZA parameter used for controlling the sparsity and convergence rate. As a result, the modified cost function of our developed algorithm is

$$I_{\text{pro}} (n) = \| \hat{w} (n + 1) - \hat{w} (n) \|^2 + \gamma_{\text{pro}} \| \hat{w} (n + 1) \|^p \quad + \lambda (d (n) - x^T (n) \hat{w} (n + 1) - \gamma).$$

The Lagrange multiplier method is also adopted to find out a minimization of cost function in (17). Then, we have

$$\frac{\partial I_{\text{pro}} (n)}{\partial \hat{w} (n + 1)} = \frac{\partial I_{\text{SM-NLMS}} (n)}{\partial \hat{w} (n + 1)} + \gamma_{\text{pro}} \frac{\partial \| \hat{w} (n + 1) \|^p}{\partial \hat{w} (n + 1)} = 0,$$

and

$$\frac{\partial I_{\text{pro}} (n)}{\partial \lambda} = 0.$$

Form (18), we get

$$\hat{w}_i (n + 1) = \hat{w}_i (n) + \frac{\lambda}{2} x_i (n) \quad \text{subject to} \quad d (n) - x_i^T (n) \hat{w}_i (n + 1) = \gamma.$$

By left-multiplying $x_i^T (n)$ on both sides of (19) and combining it with equation (20), we have

$$\frac{\lambda}{2} x_i^T (n) x_i (n) \quad \text{subject to} \quad \frac{\gamma_{\text{pro}}}{2} x_i^T (n) x_i (n) - \gamma = \frac{x_i^T (n) x_i (n)}{2 | \hat{w}_i (n + 1) |^{1-p}} \left[ \frac{\gamma_{\text{pro}}}{2} | \hat{w}_i (n + 1) |^{1-p} \right].$$

Then, substituting (21) into (19), we can get

$$\hat{w}_i (n + 1) = \hat{w}_i (n) + \frac{x_i (n)}{x_i^T (n) x_i (n)} [\gamma (n) - \gamma] \quad \text{subject to} \quad d (n) - x_i^T (n) \hat{w}_i (n + 1) = \gamma.$$

In order to prevent dividing by zero, we add a very small constant $\delta > 0$ into (23), then, the updating equation of our developed algorithm is

$$\hat{w}_i (n + 1) = \hat{w}_i (n) + \frac{x_i (n)}{x_i^T (n) x_i (n) + \delta} [\gamma (n) - \gamma] \quad \text{subject to} \quad d (n) - x_i^T (n) \hat{w}_i (n + 1) = \gamma.$$

However, we note that the ZA term in the update equation will inevitably result in error for obtaining accurate sparse channel estimation. Fortunately, we note that the parameter $p$ is a variable value according to our previous definition. Thus, we can redefine the $l_p$-norm penalty further, which is illustrated as

$$\| \hat{w} (n) \|^p_{p,N} = \sum_{i=1}^N | \hat{w}_i (n) |^p, \quad 0 \leq p \leq 1.$$ (24)

By considering the definition of (24), (23) is updated to be

$$\hat{w}_i (n + 1) = \hat{w}_i (n) + \frac{x_i (n)}{x_i^T (n) x_i (n) + \delta} [\gamma (n) - \gamma] \quad \text{subject to} \quad d (n) - x_i^T (n) \hat{w}_i (n + 1) = \gamma.$$

From (25), we can find that the zero attractor term suggests that the value of $\hat{w}_i (n + 1)$ can be divided into small and large groups. Herein, we define an expectation value

$$h (n) = E [ | \hat{w}_i (n) | ], \quad \forall 0 \leq i < N.$$ (26)

For the large group, we aim to minimize $[ p_i | | \hat{w}_i (n) |^{p_i} |]$ 0, which is subjected to $\hat{w}_i (n) > h (n)$. As for the small group, we use $p_i = 1$ to balance the solution [18]. Thereby, $p_i$ is assigned to 0 or 1 for $\hat{w}_i (n) > h (n)$ or $\hat{w}_i (n) < h (n)$, respectively. Till now, the proposed $p$-norm is separated into $l_p$-norm and $l_1$-norm by considering a mean of the channel taps. Therefore, the proposed algorithm mentioned above can be regarded as $l_p$-norm and $l_1$-norm penalty SM-NLMS for the large and small group, respectively. Then, the updating equation of (25) is changed to be

$$\hat{w}_i (n + 1) = \hat{w}_i (n) + \frac{x_i (n)}{x_i^T (n) x_i (n) + \delta} [\gamma (n) - \gamma] \quad \text{subject to} \quad d (n) - x_i^T (n) \hat{w}_i (n + 1) = \gamma.$$

In fact, we can use a matrix

$$F = \begin{bmatrix} f_0 & 0 & \cdots & 0 \\ 0 & f_1 & 0 & 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & f_{N-1} \end{bmatrix}_{N \times N}$$ (30)
to define all \( f_i \) in each iteration of (29). Then, the proposed algorithm is also written as

\[
\hat{w}(n + 1) = \hat{w}(n) + \frac{x(n)}{x^T(n)x(n) + \delta}ae(n) + \rho_{\text{pro}}F\text{sgn}[\hat{w}(n)]G(n).
\]

(31)

As we know, the proposed algorithm is complex for calculating \( G(n) \). We can use a simple method to find out the solution of (16) by considering (24) and the separation of the \( l_0 \)-norm and \( l_1 \)-norm. Then, the L0L1SM-NLMS algorithm solves an unconstrained optimization problem

\[
\min \frac{1}{2} \|\hat{w}(n + 1) - \hat{w}(n)\|^2 + \frac{1}{2} \|e(n) - \gamma\|^2_{S(n)} + \gamma_{\text{L0L1}}\|\hat{w}(n + 1)\|_p^p.
\]

(32)

Here, \( S(n) = [x^T(n)x(n)]^{-1} \), and we have \( \|e(n) - \gamma\|^2_{S(n)} = \|e(n) - \gamma\|_S^2(e(n) - \gamma) \). Then, the updating equation of our L0L1SM-NLMS is obtained

\[
\hat{w}(n + 1) = \hat{w}(n) + \frac{x(n)}{x^T(n)x(n) + \delta}ae(n) - \rho_{\text{L0L1}}F\text{sgn}[\hat{w}(n)].
\]

(33)

The developed L0L1SM-NLMS has a last zero attractor term, which is to exert \( l_0 \)-norm or \( l_1 \)-norm penalties on the channel taps according to sparse channel property in practical application. The \( l_0 \)-norm or \( l_1 \)-norm penalty is assigned via the F matrix to the separated large and small groups. Thus, our L0L1SM-NLMS algorithm can be used to enhance the convergence rate and can reduce the estimation bias for estimating sparse channels. The developed L0L1SM-NLMS algorithm synthesizes the advantages of \( l_0l_1 \)-norm and \( l_1 \)-norm penalties and serves as a dynamic adaptive norm constraint to give an \( l_0 \)-norm penalty on the large channel taps and to exert an \( l_1 \)-norm on the small channel taps. Similar to the RZASM-NLMS, the proposed L0L1SM-NLMS algorithm can be further enhanced by introducing a reweighing factor into (33) to achieve a reweighing L0L1SM-NLMS algorithm (RL0L1SM-NLMS) algorithm. As a result, the related updating equation of RL0L1SM-NLMS is

\[
\tilde{w}(n + 1) = \tilde{w}(n) + \frac{x(n)}{x^T(n)x(n) + \delta}ae(n) - \rho_{\text{RL0L1}}\tilde{f}_i\text{sgn}[\tilde{w}(n)]\frac{1}{1 + \xi|\tilde{w}(n)|}.
\]

(34)

where \( \rho_{\text{RL0L1}} \) is a ZA parameter and \( \xi \) is a positive parameter to adjust the reweighing strength. By considering all the channel taps and the training signals, (34) is changed to be

\[
\tilde{w}(n + 1) = \tilde{w}(n) + \frac{x(n)}{x^T(n)x(n) + \delta}ae(n) - \rho_{\text{RL0L1}}F\text{sgn}[\tilde{w}(n)]\frac{1}{1 + \xi|\tilde{w}(n)|}.
\]

Here, the matrix F is the same as that in (33). We observe that our RL0L1SM-NLMS algorithm has a reweighing factor \( 1/(1 + \xi|\tilde{w}(n)|) \) in its update equation in comparison with L0L1SM-NLMS algorithm. Thus, we can use the reweighing factor to adjust the zero attractor ability by properly choosing parameter \( \xi \). A suitable \( \xi \) can create an attracting constraint on the small or the large grouped channel coefficients to effectively adjust the \( l_0 \)-norm or \( l_1 \)-norm constraints, respectively.

4. Simulation Results

Here, we give an investigation on the behavior of our presented L0L1SM-NLMS and RL0L1SM-NLMS algorithms. Exactly, the convergence properties and steady-state misalignment of our developed L0L1SM- and RL0L1SM-NLMS algorithms are accessed via sparse channels with different sparse level \( K \). Furthermore, we also give the effects of the ZA parameters. The performance is obtained by the use of computer simulation and the behaviors of the L0L1SM- and RL0L1SM-NLMS algorithms are compared with classic LMS, NLMS, and SM-NLMS algorithms and popular ZA-LMS, RZA-LMS, and RZA-, RZASA-, and RZASM-NLMS. We adopt the MSE criterion to evaluate the performance of sparse channel estimators. Herein, the MSE is

\[
\text{MSE} = E\left\{\|w(n) - \hat{w}(n)\|^2\right\}.
\]

In all the experiments, \( x(n) \) denotes a Gaussian random signal which is independent of \( v(n) \). The power of \( x(n) \) and white noise is \( 1 \) and \( 1 \times 10^{-2} \), respectively. The signal-to-noise ratio (SNR) is set to be \( 20 \) dB in the experiments. In the first experiment, we consider a channel whose length is set to be \( N = 16 \), and only one nonzero tap \( (K = 1) \) is set to find out the effects of \( \rho_{\text{L0L1}} \) and \( \rho_{\text{RL0L1}} \) on our developed L0L1SM- and RL0L1SM-NLMS algorithms, respectively. Here, the simulation parameters are \( \delta = 0.01 \) and \( \gamma = \sqrt{2\delta} \) in L0L1SM-NLMS algorithm. Figure 1 gives the effects of \( \rho_{\text{L0L1}} \) on our L0L1SM-NLMS algorithm. As we can see form Figure 1, there is a large MSE for \( \rho_{\text{L0L1}} = 1 \times 10^{-1} \). When \( \rho_{\text{L0L1}} \) decreases from \( 5 \times 10^{-2} \) to \( 1 \times 10^{-3} \), the MSE of our developed L0L1SM-NLMS algorithm is reduced, indicating that our L0L1SM-NLMS algorithm achieves a low estimation misalignment. If \( \rho_{\text{L0L1}} \) continues to decrease, the MSE of our L0L1SM-NLMS algorithm is getting larger. Next, the effect of \( \rho_{\text{RL0L1}} \) is shown in Figure 2 with \( \xi = 0.05 \) and \( \gamma = \sqrt{2\delta} \). It is clear to see that the MSE gradually decreases with \( \rho_{\text{RL0L1}} \) reducing from \( 1 \times 10^{-1} \) to \( 5 \times 10^{-5} \). When \( \rho_{\text{RL0L1}} \) ranges from \( 1 \times 10^{-3} \) to \( 5 \times 10^{-5} \), the MSE is getting larger with the decrement of \( \rho_{\text{RL0L1}} \). Thus, \( \rho_{\text{L0L1}} \) and \( \rho_{\text{RL0L1}} \) should be properly selected to well improve the behavior of our presented L0L1SM- and RL0L1SM-NLMS algorithms.

Then, we will verify estimation behaviors of the L0L1SM- and RL0L1SM-NLMS algorithms over a sparse channel with different sparsity levels and their estimation behaviors are compared with LMS, NLMS, SM-NLMS, and their sparse forms. The convergence of our presented L0L1SM- and RL0L1SM-NLMS algorithms compared with NLMS algorithms is shown in Figure 3. We observe that both L0L1SM- and RL0L1SM-NLMS algorithms converge faster than those
of the previously discussed LMS and NLMS algorithms. It is worth pointing out that our developed RL0L1SM-NLMS algorithm provides the fastest convergence speed. The estimation performance in terms of the steady-state misalignment is shown in Figures 4, 5, and 6 for $K = 2$, $K = 4$, and $K = 8$, respectively. In the experiments, the simulation parameters are $\mu_{NLMS} = 0.5$ for NLMS algorithm and its sparse forms, and $\rho_{ZASM}$, $\rho_{ZANLMS}$, and $\rho_{RZASM}$ are the ZA parameters for the already reported ZASM-, ZA-, and RZASM-NLMS algorithms, respectively. The steady-state misalignment of our L0L1SM- and RL0L1SM-NLMS gives lower estimation error floors compared to the traditional
algorithm can provide lowest estimation misalignment compared to the recent reported ZASM-NLMS and RZASM-NLMS algorithm when $K$ equals 2. When $K$ increases from 4 to 8, the steady-state misalignment becomes higher in comparison with that of $K = 2$. However, the steady-state error is better than those sparse SM-NLMS and NLMS algorithms and their related sparse algorithms realized by using ZA techniques.

The steady-state behaviors of our L0L1SM- and RL0L1SM-NLMS algorithms are fully verified and well compared with the existing LMS algorithms, including LMS, ZA-LMS, RZA-LMS, ZA-NLMS, and RZA-NLMS. Here, the parameters of our L0L1SM- and RL0L1SM-NLMS algorithms are the same as those in the last experiment, while other extra simulated parameters are listed as follows: $\mu_{LMS} = 0.05$, $\rho_{ZA-LMS} = 5 \times 10^{-4}$, and $\rho_{RZA-LMS} = 9 \times 10^{-4}$, where $\mu_{LMS}$ is a step size for LMS algorithm and $\rho_{ZA-LMS}$ and $\rho_{RZA-LMS}$ are the ZA parameters for ZA- and RZA-LMS algorithms, respectively. The estimation behaviors for $K = 2$, $K = 4$, and $K = 8$ are described in Figures 7, 8, and 9, respectively. It is found that our RL0L1SM-NLMS algorithm achieves fast convergence and possesses the lowest estimation misalignment. Additionally, the proposed L0L1SM-NLMS algorithm is also better than the mentioned algorithms in terms of MSE. When the number of the dominant coefficients increases to 8, the steady-state misalignment of our developed algorithms increased. However, our developed algorithms still yield excellent steady-state behavior, implying that our presented L0L1SM- and RL0L1SM-NLMS algorithms are more powerful than the popular LMS and its sparse versions. In addition, the estimation behaviors of our constructed L0L1SM- and RL0L1SM-NLMS algorithms are discussed in Figure 10. We can see that the estimation misalignment of RL0L1SM-NLMS algorithm is better than the L0L1SM-NLMS algorithm for any sparsity level $K$, indicating the proposed RL0L1SM-NLMS algorithm is more robust.

Finally, the estimation behavior of our L0L1SM- and RL0L1SM-NLMS algorithms is studied for estimating an echo channel which is illustrated in Figure 11 as an example. There are 16 nonzero taps ($K = 16$) within the echo channel and its length is 256 ($N = 256$). In this experiment, the sparsity is illustrated as $\zeta_1(w) = (N/(N - \sqrt{N})) (1 - \|w\|_1/\sqrt{N}\|w\|_2)$ [28, 35]. The SNR is 30 dB and the used parameters in this experiment are $\mu_{NLMS} = \mu_{ZA-NLMS} = 0.5$, $\rho_{ZA-NLMS} = 3 \times 10^{-5}$, $\rho_{RZA} = 5 \times 10^{-6}$, $\rho_{RZASM} = 6 \times 10^{-6}$, $\rho_{L0L1} = 7 \times 10^{-5}$, $\rho_{R0L1} = 1 \times 10^{-5}$, and the other parameters are the same as the above experiments. We use $\zeta_1(w) = 0.8222$ and $\zeta_1(w) = 0.7362$ for the first 10000 and second 10000 iterations, respectively. Herein, only the NLMS and its sparse forms are used for comparison since the NLMS algorithm is better than the LMS algorithm [22]. The computer simulation result is given in Figure 12. As is described in Figure 12, the steady-state misalignment of our developed L0L1SM- and RL0L1SM-NLMS algorithms is superior to the previously developed RZASM-NLMS algorithm. After the first 10000 iterations, the steady-state error of our algorithms is getting larger. However, their estimation behaviors are still better than mentioned channel estimation algorithms. Therefore,
5. Conclusion

A L0L1SM-NLMS algorithm and a RL0L1SM-NLMS algorithm have been proposed and their derivations have been introduced in detail. The proposed L0L1SM-NLMS algorithm was realized by integrating a $p$-norm and separating the $p$-norm into $l_0$ and $l_1$ norm via a mean of the channel taps.

A conclusion is given that our developed L0L1SM- and RL0L1SM-NLMS algorithms have an excellent performance based on the evaluation criterion of convergence and estimation behavior for sparse adaptive channel estimation.
and the RL0L1SM-NLMS algorithm was implemented by using a reweighting factor in the L0L1SM-NLMS algorithm to enhance the ZA strength. Both the proposed algorithms can provide a ZA-like performance on the small channel coefficients and give rise to an \(l_0\)-norm penalized ZA on the large channel taps. The proposed L0L1SM- and RL0L1SM-NLMS algorithms have been discussed via a sparse channel with different sparsity levels and a designated echo channel. The computer simulations obtained from these channels give a conclusion that our RL0L1SM-NLMS algorithm has the best behavior with respect to the convergence and steady-state channel estimation behavior. Furthermore, both the proposed L0L1SM- and RL0L1SM-NLMS algorithms provide better estimation behaviors than the traditional LMS, NLMS, and SM-NLMS and the previously reported popular sparse algorithms.

**Competing Interests**

The authors declare that there is no conflict of interests regarding the publication of this article.

**Acknowledgments**

This work was partially supported by the National Key Research and Development Program of China-Government Corporation Special Program (2016YFE011100), the Science and Technology Innovative Talents Foundation of Harbin (2016RAXXJ044), the Foundational Research Funds for the Central Universities (HEUCFD1433, HEUCF160815), and Projects for the Selected Returned Overseas Chinese Scholars of MOHRSS and Heilongjiang Province of China.

**References**


