IET Communications

Reweighted lp Constraint LMS-Based Adaptive Sparse Channel Estimation for Cooperative Communication System

COM-2018-6186.R2 | Research Article

Submitted on: 29-05-2019

Submitted by: Aihua Zhang, Pengcheng Liu, Bing Ning, Qiyu Zhou

Keywords: LMS, COOPERATIVE COMMUNICATION, SPARSE MULTIPATH CHANNELS, CHANNEL ESTIAMTION

Reweighted l_p Constraint LMS-Based Adaptive Sparse Channel Estimation for Cooperative Communication System

Aihua Zhang[1*], Pengcheng Liu[2], Bing Ning[1], Qiyu Zhou[1]

1.School of Electronic and Information Engineering, Zhongyuan University of Technology, Zhengzhou, 450007, China;

2 Cardiff School of Technologies, Cardiff Metropolitan University, Cardiff CF5 2YB, UK.

[*]corresponding author: zhah@zut.edu.cn

Abstract: This paper studies the issue of sparsity adaptive channel reconstruction in time-varying cooperative communication networks through the amplify-and-forward transmission scheme. A new sparsity adaptive system identification method is proposed, namely reweighted l_p norm ($0 < p < 1$) penalized least mean square (LMS) algorithm. The main idea of the algorithm is to add a l_p norm penalty of sparsity into the cost function of the LMS algorithm. By doing so, the weight factor becomes a balance parameter of the associated l_p norm adaptive sparse system identification. Subsequently, the steady state of the coefficient misalignment vector is derived theoretically, with a performance upper bounds provided which serve as a sufficient condition for the LMS channel estimation of the precise reweighted l_p norm. With the upper bounds, we prove that the l_p ($0 < p < 1$) norm sparsity inducing cost function is superior to the reweighted l_1 norm. An optimal selection of p for the l_p norm problem is studied to recover various d sparse channel vectors. Several experiments verify that the simulation results agree well with the theoretical analysis, and thus demonstrate that the proposed algorithm has a better convergence speed and better steady state behavior than other LMS algorithms.

1. Introduction

Cooperative communication has been widely studied recently in wireless networks because of its significant performance in enhancing the transmission capacity and exploiting spatial diversity to against the influence of path loss and channel fading [1]. In cooperative communication systems, the accurate channel impulse response (CIR) is needed for equalization, coherent signal detection, and so on, and it can also improve the communication quality of service in 5G wireless communication systems, especially for the dynamically changing channel and its sparsity. Therefore, the issue of accurately estimate the channel state information in dynamic cooperative relay channel systems becomes significant and challenging.

In cooperative communication systems, the multipath tap-weights are spread widely in time with only a few significant components, and the impulse response of multipath wireless channel contains only a small fraction of nonzero coefficients, which means the cooperative channel has sparse structure. As such, the cooperative relay multipath wireless channel is the cooperative relay multipath wireless channel is characterized as a fast time-varying and sparse feature. By utilizing and exploiting the inherent sparsity of cooperative channel impulse response, the channel estimation performance can be improved. Currently, there has been a growing interest in sparse channel estimation, and advanced channel estimation algorithms have been developed such as compressed sensing (CS) algorithms and sparse adaptive filtering (SAF) algorithms [2–6] and so forth.

Sparse channel estimation methods mainly include: optimization methods, thresholding methods and greedy methods. Classic algorithms include basis pursuit (BP) algorithm, orthogonal matching pursuit (OMP) method and iterative thresholding algorithms [7-9]. Unfortunately, these algorithms are not applicable for sparse channel estimation in fast time-varying environments. In [3], the sparsity adaptive expectation maximization (SAEM) used expectation maximization algorithm (EM) and Kalman filter which can utilize channel sparsity well and trace the true support set of time-varying channel [3]. However, SAEM has high computational complexity.

Accordingly, the LMS-based sparse adaptive filtering or recursive least squares (RLS) algorithms are developed attribute to their simplicity in application [10–13]. In addition, a class of novel sparse adaptive algorithms has emerged based on regularized LMS algorithms, where the sparsity penalty induced strategy is used by exerting various sparsity penalty terms into the instantaneous error of a traditional adaptive filtering algorithm. The sparsity constraint can be l_1 norm [10], reweighted l_1 norm [12], l_0 norm [13], and non-convex sparsity penalty. These algorithms have good performance on faster convergence rate and smaller mean square error (MSE) comparing with the traditional adaptive filtering method, such as zero-point attraction Least Mean Square algorithm (ZA-LMS) [10], reweighted zero attracting LMS (RZA-LMS) [11] and so on. ZA-LMS uses a l_1 norm penalty in the cost function of the traditional LMS algorithm, where l_1 norm acts as a zero-point attracted term to modify the parameter vector update equation. RZA-LMS introduces the log-sum penalty and its performance is similar to the l_0 norm algorithm. Y. Gu proposed the l_0 norm Constraint LMS Algorithm [13], by exerting the l_0 norm penalty into the cost function of the LMS algorithm. The l_0 norm, a more accurate measure of sparsity, is defined as the number of non-zero elements in the unknown system vector. Similarly, as proposed in [14], ZA-RLS-I and ZA-RLS-II added l_1 norm penalty and approximated l_1 norm of the parameter vector penalty term instead of an adaptively weighted l_2 norm penalty to cost function of the RLS algorithm. The ZA-RLS algorithms achieve better performance than the other LS algorithms, however, their MSEs are not as good as sparse LMS algorithms.

Recently, the non-convex methods have received tremendous attentions in solving the problem of sparse recovery [15-18]. Furthermore, some studies have presented that the non-convex penalties might induce better sparsity than the convex penalties. In addition, the local and global optimality of l_p minimization for sparse recovery can be guaranteed even under weaker conditions in comparison with the convex l_1 minimization when the penalty approaches the l_p norm [17]. In this work, we study the fast identification of sparse cascaded channel by using the framework of an adaptive filter. In order to explore the sparse features of the cooperative relay communication system, we propose a new sparse aware LMS algorithm for relay channel reconstruction. The expectation of the misalignment vector is derived and discussed under different algorithm parameters and system sparsity. Simulation studies are conducted to verify the high robustness, low computational cost and easy implementation of the proposed algorithm.

This paper is organized as follows. The amplify-and-forward (AF)-based cooperative relay channel model is described briefly in Section 2. Then in Section 3, we introduce the reweighted l_p norm constraint LMS and derive the expectation of the misalignment vector and provide the steady-state analysis of the proposed algorithm. Numerical simulations and rigorous analysis are presented in Section 4 to demonstrate the effectiveness prove the theoretical analysis. Finally, the conclusion is given in section 5.

2. System model and LMS algorithm

2.1 Cooperative Rely Channel Model

Consider an amplified model of cooperation relay network with a source node $\mathbf{T_1}$, a destination node $\mathbf{T_2}$, and one relay node $\mathbf{R}$. It is assumed that all the terminals are equipped with only one antenna and work in the half-duplex mode. When node $\mathbf{T_2}$ is beyond the communication range of node $\mathbf{T_1}$ duo to remote distance or shielding affection, then all signals sent by the source $\mathbf{T_1}$ need to be forwarded to destination $\mathbf{T_2}$ by relay node. Denote $\boldsymbol{g} = [g_0, g_1, \cdots g_{L_g-1}]$ as the baseband channel between $\mathbf{T_1}$ and $\mathbf{R}$. And $\boldsymbol{k} = [k_0, k_1, \cdots k_{L_k-1}]$ is the channel vector between relay node $\mathbf{R}$ and destination $\mathbf{T_2}$. Since $\mathbf{T_1}$ and $\mathbf{T_2}$ are separated from each other, $\boldsymbol{g}$ and $\boldsymbol{k}$ are considered independent. The taps of all these two channels are assumed as zero-mean circularly symmetric complex Gaussian random variables, i.e., $g_i \sim \mathcal{CN}(0, \sigma_{g,i}^2)$, $k_i \sim \mathcal{CN}(0, \sigma_{k,i}^2)$. Moreover, the source and relay are assumed to have average power constraints, which are denoted by P_1 and P_R, respectively.

There are two stages in amplified relay transmission system, and it takes two time slots to achieve the cooperative multiple access. In the first time slot, the source node $\mathbf{T_1}$ sends signals $\boldsymbol{x}$ and the relay node $\mathbf{R}$ receives as

$$\boldsymbol{r} = \boldsymbol{g}\boldsymbol{x} + \boldsymbol{n}_r, \quad (1)$$

where $\boldsymbol{n}_r$ is the additive white Gaussian noise with variance σ_r^2.

In the second time slot, the relay node amplifies and transmits the received data to the destination node $\mathbf{T_2}$, and $\mathbf{T_2}$ receives as

$$\boldsymbol{y}(n) = \underbrace{\alpha\boldsymbol{k}\boldsymbol{g}}_{\boldsymbol{h}}\boldsymbol{x} + \underbrace{\alpha\boldsymbol{k}\boldsymbol{n}_r + \boldsymbol{n}_1}_{\boldsymbol{n}}$$

$$= \boldsymbol{hx} + \boldsymbol{n}, \tag{2}$$

where $\boldsymbol{h}$ (with the length of $L = L_g + L_k - 1$) is the cascaded channel that is the convolution between $\boldsymbol{g}$ and $\boldsymbol{k}$, $\boldsymbol{n}_1$ represents the noise at $\mathbf{T}_2$ with variance σ_1^2, $\boldsymbol{n}$ denotes the overall noise. $\alpha = \sqrt{P_{\mathrm{R}}/[P_1\sigma_g^2 + \sigma_r^2]}$ and $\sigma_g^2 = \sum_{i=0}^{L_g-1}\sigma_{g,i}^2$.

2.2 Standard LMS

In the AF relay cooperative communication system, the unknown cascaded channel coefficients at time instant n are $\boldsymbol{h} = [h_0, h_1, \cdots, h_{L-1}]^T$. The system's input data vector from $\mathbf{T}_1$ is expressed as $\boldsymbol{x} = [x_n, x_{n-1}, \cdots, x_{n-L+1}]^T$ and it is assumed to be independent Gaussian input. As shown in Fig. 1, we consider sparse adaptive channel estimation in a relay-based cooperation communication system.

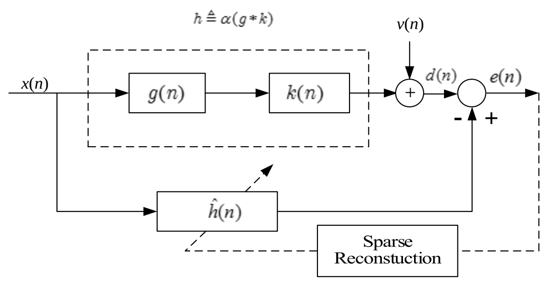


Fig. 1 Channel model and adaptive channel estimation algorithm

In Fig. 1, the desired output signal $\boldsymbol{d}(n)$ accord to $\boldsymbol{y}(n)$, denoted as

$$\boldsymbol{d}(n) = \boldsymbol{h}^T(n)\boldsymbol{x}(n) + \boldsymbol{v}(n), \tag{3}$$

where $\boldsymbol{v}(n)$ is the noise signal. The estimated error $\boldsymbol{e}(n)$ is the instantaneous error between the output signal of the unknown system and the output from the adaptive filter, which can be written as

$$\boldsymbol{e}(n) = \boldsymbol{d}(n) - \widehat{\boldsymbol{h}}^T(n)\boldsymbol{x}(n), \tag{4}$$

where $\widehat{\boldsymbol{h}} = [\widehat{\boldsymbol{h}}_0, \widehat{\boldsymbol{h}}_1, \cdots \widehat{\boldsymbol{h}}_{L-1}]^T$ is defined as an adaptive tap-weights vector. The noise vector $\boldsymbol{v}(n)$ follows i.i.d. zero mean and δ_n^2 variance white Gaussian distribution. It is assumed that the adaptive tap-weights vector $\boldsymbol{h}(n)$, input signal $\boldsymbol{x}(n)$ and additive noise signal $\boldsymbol{v}(n)$ are mutually independent.

According to the standard LMS framework, the cost function is defined as $\xi(n) = 0.5|\boldsymbol{e}(n)|^2$. The recursive equation of the filter coefficient vector can be derived as

$$\widehat{\boldsymbol{h}}(n+1) = \widehat{\boldsymbol{h}}(n) - \mu\frac{\partial\xi(n)}{\partial\widehat{\boldsymbol{h}}(n)} = \widehat{\boldsymbol{h}}(n) + \mu\boldsymbol{e}(n)\boldsymbol{x}(n), \tag{5}$$

where μ is the step size parameter which satisfies $\mu \in (0, 1/\lambda_{max})$ and adjusts the convergence rate and the steady performance behavior of the LMS algorithm[4]. λ_{max} is the maximum eigenvalue of P_x, here $P_x = E[\boldsymbol{x}(n)\boldsymbol{x}^T(n)]$, which denotes the covariance matrix of the input vector $\boldsymbol{x}(n)$.

3. The proposed channel estimation algorithm

This work focuses on fast unknown channel identification of the cooperation system via sparse constraint adaptive filter. The impulse response of a sparse cooperative system consists of few nonzero coefficients, most of the coefficients in the channel representation vector of $\boldsymbol{h}(n)$ in time delay domain should be zeroes or small values. In order to improve the performance of sparsity adaptive channel estimation, we propose a novel sparsity-aware system identification method with a new cost function.

3.1 Reweighted l_p Norm Penalized LMS

For the sake of exploiting the sparse structure of cooperative relay communication system, we propose an idea of introducing reweighted l_p norm of channel impulse response as a sparsity penalty into the cost function of traditional LMS algorithm. This method can accelerate and enhance the performance of sparse cascade channel estimation.

Motivated by the research findings that non-convex penalties might induce better sparsity than the convex penalties [18], we apply reweighted l_p $(0 < p < 1)$ norm to measure the sparsity of channel vector.

$$\|\widehat{\boldsymbol{h}}(n)\|_p = (\textstyle\sum_i|\widehat{\boldsymbol{h}}(n)|^p)^{\frac{1}{p}}, \tag{6}$$

where $\|\cdot\|_p$ stands for the l_p norm of the channel vector.

When $0 < p < 1$

$$\lim_{p\to 1}\|\widehat{\boldsymbol{h}}(n)\|_p = \|\widehat{\boldsymbol{h}}(n)\|_1 = \textstyle\sum_{i=1}^{L}|\widehat{\boldsymbol{h}}(i)|.$$

$$\lim_{p\to 0}\|\widehat{\boldsymbol{h}}(n)\|_p = \|\widehat{\boldsymbol{h}}(n)\|_0.$$

In order to learn the time-varying channel state information by using the prior sparsity, we apply a new cost function, which combines the instantaneous channel estimation square error and the l_p $(0 < p < 1)$ norm penalty term of the coefficient vector, defined as

$$\xi_p(n) = \frac{1}{2}|\boldsymbol{e}(n)|^2 + \gamma_p\|\widehat{\boldsymbol{h}}(n)\|_p, \tag{7}$$

where γ_p can be selected using a positive factor to balance the mean square error and adjust the penalty of l_p norm.

The gradient of $\xi_p(n)$ is

$$\frac{\partial \xi_p(n)}{\partial \hat{\boldsymbol{h}}(n)} = -\boldsymbol{e}(n)\boldsymbol{x}(n) + \gamma_p \frac{\left(\|\hat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)} sgn(\hat{\boldsymbol{h}}(n))}{|\hat{\boldsymbol{h}}(n)|^{(1-p)}}. \quad (8)$$

On the right-hand side of equation (8), the second term is the gradient of l_p norm.

Following the gradient descent algorithm, we use the gradient of l_p norm as the zero attracting term. The reweighted l_p norm LMS algorithm updates its coefficients by (9)

$$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu \boldsymbol{e}(n)\boldsymbol{x}(n) - \rho_p \frac{\left(\|\hat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)} \mathrm{sgn}(\hat{\boldsymbol{h}}(n))}{\varepsilon_p + |\hat{\boldsymbol{h}}(n)|^{(1-p)}}, \quad (9)$$

where $\rho_p = \alpha\mu\gamma_p$ with $\alpha\ (0 < \alpha < 1)$ is a reweighted parameter adopted to adjust zero-point attraction ability of the l_p norm penalty. We set parameter $\varepsilon_p > 0$ for providing stability and ensuring that a zero value in $\hat{h}(n)$ does not prohibit a non-zero estimate strictly [6]. Additionally, ε_p should be set as a small positive value or smaller than the expected nonzero magnitudes of $\hat{h}(n)$. In consequence, the reasonably robust channel estimation process tends to be dependent on the choice of ε_p.

The procedure of the reweighted l_p norm LMS algorithm is listed in Table 1, where 0_L is a zero vector of size L.

Table 1 The proposed reweighted l_p norm LMS algorithm

Require: $\mu > 0, \gamma > 0, \alpha > 0,\ \varepsilon_p > 0\ ,\ 0 < p < 1,$
$\boldsymbol{g}$ and $\boldsymbol{k}$ as zeroes or small random vectors;

1. **Initialize** $\hat{h}(0) = 0_L, e(0) = 0,\ i$=1;
2. while i< N
3. Calculate the error through equation (4);
4. Update the gradient of $\xi_p(n)$ according to equation (7);
5. Update the reweighted zero attractor by multiplying α

$$\alpha \times \mu\gamma_p \times \left(\|\hat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)} \times \mathrm{sgn}(\hat{\boldsymbol{h}}(n)/[\varepsilon_p + |\hat{\boldsymbol{h}}(n)|^{(1-p)}].$$

6. Update the tap-weight vector according to equation (9)
7. $i = i + 1$;
8. end while

Reweighted l_1 norm penalty: Another method of exploring the sparsity of the wireless communication channel is to use a reweighted l_1 norm constraint in the cost function of SAF algorithms [12]. The reweighted l_1 norm needs to be minimized, which approximates the l_0 pseudo norm better than the l_1 norm. The reweighted l_1 norm penalized LMS algorithm considers a l_1 norm penalty term to channel impulse response vector. The reweighted l_1 norm cost function in [12] is derived as

$$\xi_1(n) = \frac{1}{2}|\boldsymbol{e}(n)|^2 + \gamma_1 \left\|\boldsymbol{s}(n)\hat{\boldsymbol{h}}(n)\right\|_1,$$

where γ_1 is the weight for adjusting the l_1 norm penalty term. And $\boldsymbol{s}(n)$ acts as a weight element that can be denoted as

$$[\boldsymbol{s}(n)]_i = \frac{1}{\varepsilon_1 + \left|[\hat{\boldsymbol{h}}(n-1)]_i\right|}, i = 1,2,\cdots,L.$$

where ε_1 should be set as a small positive value and $[\cdot]_i$ is the ith entry of the estimated channel coefficient vector. The gradient of $\xi_1(n)$ can be written as

$$\frac{\partial \xi_1(n)}{\partial \hat{\boldsymbol{h}}(n)} = -\boldsymbol{e}(n)\boldsymbol{x}(n) + \gamma_1 \frac{\mathrm{sgn}(\hat{\boldsymbol{h}}(n))}{\varepsilon_1 + |\hat{\boldsymbol{h}}(n-1)|}.$$

The resulting update is

$$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu\boldsymbol{e}(n)\boldsymbol{x}(n) - \rho_1 \frac{\mathrm{sgn}(\hat{\boldsymbol{h}}(n))}{\varepsilon_1 + |\hat{\boldsymbol{h}}(n-1)|},$$

where $\rho_1 = \mu\gamma_1$.

3.2. Computational Complexity

Computational complexity of the proposed reweighted l_p Constraint LMS algorithm and some sparsity-aware LMS algorithms are compared in Table 2, in terms of arithmetic operations and comparisons. As shown in Table 2, the proposed algorithm has lower computation complexity than l_0-LMS [13]. And the amount of computations of the proposed algorithm is similar to that of RZA-LMS and the reweighted l_1 norm penalized LMS.

Table 2 Computational complex of different algorithms

Algorithm	Computational Complexity
LMS	(2L) Add+(2L+1) Multiply
RZA-LMS	(4L) Add+(5L+1) Multiply
l_0-LMS	(4L) Add+(5L+1) Multiply+(L) Comp
Reweighted l_1-LMS	(4L) Add+(5L+1) Multiply
Proposed algorithm	(4L) Add+(5L+1) Multiply

3.3 Performance Analysis

In this section, we derive the theoretical steady-state of the coefficient misalignment and provide a mean square error convergence analysis of the new sparsity adaptive channel estimation algorithm. Then a performance upper bounds is drawn as a sufficient condition for the precise reweighted l_p norm LMS channel estimation.

A. Mean Performance

Assuming an i.i.d. zero-mean Gaussian input signal $\boldsymbol{x}(n)$ and a zero mean white noise. We define $\boldsymbol{r}(n)=\widehat{\boldsymbol{h}}(n)-\boldsymbol{h}(n)$ as the filter misalignment vector. The recursion formula of the misalignment vector can be written as

$$\boldsymbol{r}_{n+1}=(\boldsymbol{I}-\mu\boldsymbol{x}_n\boldsymbol{x}_n^T)r_n+\mu\boldsymbol{v}_n\boldsymbol{x}_n-\rho_p f(\widehat{\boldsymbol{h}}_n), \quad (10)$$

where $f(\widehat{\boldsymbol{h}}_n)$ is defined as

$$f(\widehat{\boldsymbol{h}}_n)=\frac{\left(\|\widehat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)}\operatorname{sgn}(\widehat{\boldsymbol{h}}(n))}{\varepsilon_p+|\widehat{\boldsymbol{h}}(n)|^{(1-p)}}. \quad (11)$$

Taking expectation, since $\boldsymbol{v}_n$ is assumed to have a zero mean and be independent with input signal $\boldsymbol{x}(n)$. Assume that $\boldsymbol{h}(n)$ is a d-sparse channel vector, we have

$$E[\boldsymbol{r}_{n+1}]=(I-\mu P_x)E[\boldsymbol{r}_n]-\rho_p E[f(\widehat{\boldsymbol{h}}_n)]. \quad (12)$$

$$E[f(\widehat{\boldsymbol{h}}_n)]=E[\frac{\left(\|\widehat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)}\operatorname{sgn}(\widehat{\boldsymbol{h}}(n))}{\varepsilon_p+|\widehat{\boldsymbol{h}}(n)|^{(1-p)}}]. \quad (13)$$

$$-\frac{\rho_p}{\mu P_x}\mathbf{1}<E[r_\infty]<\frac{\rho_p}{\mu P_x}\mathbf{1}. \quad (14)$$

Further derivation, we obtain

$$-E[r_\infty]<\frac{\rho_p}{\mu P_x}\mathbf{1}\le\frac{\rho_p(\sqrt[p]{d})^{1-p}}{\mu P_x[\varepsilon_p+|\widehat{\boldsymbol{h}}(n)|^{(1-p)}]}\mathbf{1}, \quad (15)$$

where d denotes the number of non-zero coefficients, which means the sparsity level of cooperative communication channel. It can be observed that $E[r_\infty]$ is bounded between $-\rho_p\mathbf{1}/(\mu P_x)$ and $\rho_p\mathbf{1}/(\mu P_x)$, where $\mathbf{1}$ is the vector with all one entries. This means that the reweighted l_p norm LMS algorithm has a stability condition for the coefficient misalignment vector convergence. From (15), we can achieve better performance by adjusting parameter ε_p by following the change of sparsity d.

The mean misalignment vector of the reweighted l_1 norm penalized LMS algorithm is bounded as

$$-\frac{\rho_1}{\mu P_x\varepsilon_r}\mathbf{1}\le E[r_\infty]\le\frac{\rho_1}{\mu P_x\varepsilon_r}\mathbf{1}. \quad (16)$$

when $\rho_1=\rho_p$, as ε_r is a very small positive value, so $\rho_p\mathbf{1}/(\mu P_x)<\rho_1\mathbf{1}/(\mu P_x\varepsilon_r)$ generally. Theoretically, the performance of reweighted l_p norm sparse aware LMS algorithm is better than the reweighted l_1 norm penalized LMS algorithm.

B. Mean Square Steady-State Performance of proposed algorithm

The mean square deviation (MSD) bounds of the proposed reweighted l_p norm constraint LMS are derived by the following theorem. In order to guarantee convergence, step-size μ should satisfy

$$0<\mu<\frac{2}{(L+2)P_x}. \quad (17)$$

The final mean square deviation of the proposed algorithm is

$$S(\infty)=\frac{2[1-\mu P_x]\gamma c(\infty)+\gamma^2\mu q(\infty)+L\mu P_v P_x}{P_x[2-(L+2)\mu P_x]}, \quad (18)$$

where $c(n)=E\left[r^T(n)\boldsymbol{f}\left(\hat{h}(n)\right)\right]$, $q(n)=\left\|\boldsymbol{f}(\hat{h}(n)\right\|_2^2$, $c(n)$ and $q(n)$ are all bounded. The proof of (17) and (18) goes in Appendix.

From (17), we can conclude that the MSD will decrease as the channel length L increases, which has been proved in the simulations. It is shown that the convergence of the proposed algorithm can be guaranteed when μ satisfies (17). When p closes to zero, steady-state MSD of our proposed algorithm will be smaller and then the steady state performance will be better than other sparse LMS algorithm.

4. Simulation results and analysis

In this section, three experiments are provided to demonstrate the estimation performance of sparse adaptive filtering algorithms. The reference algorithms simulated for comparison with the proposed algorithm include Standard LMS [4], ZA-LMS [10], RZA-LMS [11], reweighted $l_1$norm penalized LMS algorithms [12], and l_0 LMS [13]. The cost functions and updated equations of the above algorithms are listed in Table 3.

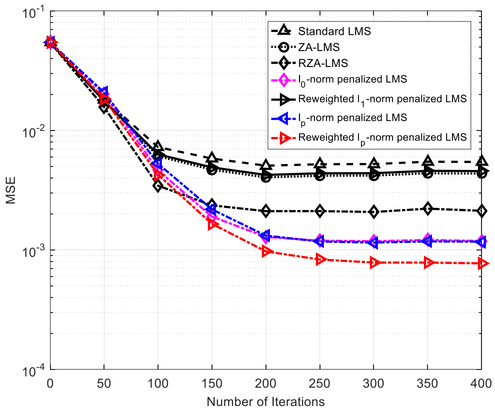
Table 3 Various LMS algorithms

Algorithm	Cost function	Update equation
LMS	$\xi(n)=\frac{1}{2}\|e(n)\|^2$	$\widehat{\boldsymbol{h}}(n+1)=\widehat{\boldsymbol{h}}(n)+\mu e(n)\boldsymbol{x}(n)$
ZA-LMS	$\xi_{ZA}(n)=\frac{1}{2}\|e(n)\|^2+\gamma_{ZA}\|\widehat{\boldsymbol{h}}(n)\|$	$\widehat{\boldsymbol{h}}(n+1)=\widehat{\boldsymbol{h}}(n)+\mu e(n)\boldsymbol{x}(n)$ $-\rho_{ZA}\operatorname{sgn}[\widehat{\boldsymbol{h}}(n)]$
RZA-LMS	$\xi_{RZA}(n)=\frac{1}{2}\|e(n)\|^2$ $+\gamma_{RZA}\sum_{l=1}^{L}\log(1+\widehat{\boldsymbol{h}}(n))$	$\widehat{\boldsymbol{h}}(n+1)=\widehat{\boldsymbol{h}}(n)+\mu e(n)\boldsymbol{x}(n)$ $-\rho_{RZA}\frac{\operatorname{sgn}[\widehat{\boldsymbol{h}}(n)]}{1+\varepsilon_{RZA}\|\widehat{\boldsymbol{h}}(n)\|}$
l_0-LMS	$\xi_0(n)=\frac{1}{2}\|e(n)\|^2+\gamma_0\|\widehat{\boldsymbol{h}}(n)\|_0$	$\widehat{\boldsymbol{h}}(n+1)=\widehat{\boldsymbol{h}}(n)+\mu e(n)\boldsymbol{x}(n)$ $-\rho_0\beta\operatorname{sgn}[\widehat{\boldsymbol{h}}(n)]e^{-\beta\|\widehat{\boldsymbol{h}}(n)\|}$
Proposed algorithm	$\xi_p(n)=\frac{1}{2}\|e(n)\|^2+\gamma_p\|\widehat{\boldsymbol{h}}(n)\|_p$	$\widehat{\boldsymbol{h}}(n+1)=\widehat{\boldsymbol{h}}(n)+\mu e(n)\boldsymbol{x}(n)$ $-\rho_p\frac{\left(\|\widehat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)}\operatorname{sgn}(\widehat{\boldsymbol{h}}(n))}{\varepsilon_p+\|\widehat{\boldsymbol{h}}(n)\|^{(1-p)}}$

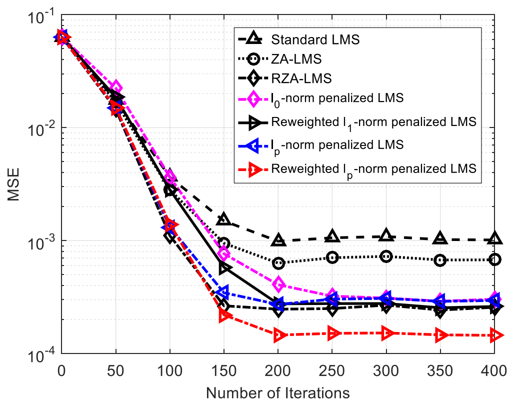
Experiment 1, we assume that the channel vectors of cooperative relay system have the same length $L_g=L_k=16$ and hence the length of convolution channel vectors is $L=L_g+L_k-1=31$. In case 1, two large coefficients of $\boldsymbol{g}$ and $\boldsymbol{k}$ are uniformly distributed and all the others are exactly zero, making the system have a sparsity level of $2/_{31}$. Four random tap-weights of $\boldsymbol{g}$ and $\boldsymbol{k}$ are nonzero in case 2.

The values of several large coefficients are chosen from Gaussian distribution with a zero mean and a unit variance.

The parameters of reweighted l_p norm LMS channel estimation algorithm and reweighted l_1 norm LMS algorithm are set to $\rho_p = \rho_1 = 5 \times 10^{-4}$ and $\varepsilon_p = \varepsilon_1 = 1.2$. In the ZA-LMS algorithm and RZA-LMS algorithm, the parameters are set as $\rho_{ZA} = \rho_{RZA} = 5 \times 10^{-4}$ and $\varepsilon_{RZA} = 10$. We set step size parameter $\mu = 0.02$ for all algorithms in this paper. We test the performance of the proposed algorithm under the low Signal-to-Noise Ratio (SNR) 10dB and high SNR 20dB respectively. The average estimation of the mean square error (MSE) between the actual and estimated channel state information are shown in Fig. 2 and Fig. 3.



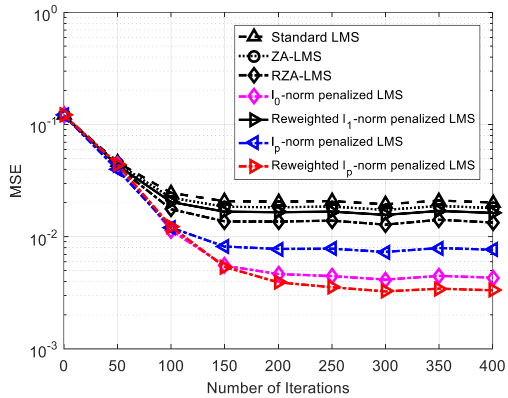
a



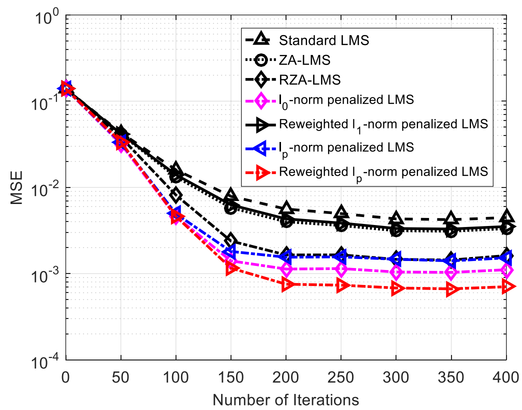
b

Fig. 2 Example 1, Case 1: Comparison of convergence rate for six different algorithms ($L = 31$, $d = 2$).
(***a***) SNR=10 dB, (**b**) SNR=20 dB

The estimated channel impulse response MSE results for sparsity case 1 are shown in Fig. 2 (a)~ (b) and for sparsity case 2 are shown in Fig. 3 (a) ~ (b). According to Figs. 2~3, when the sparsity of the channel increases, the convergence performance of the sparsity-aware parameter estimation algorithms degrades accordingly. By examining the convergence lines, we can conclude that, in general, the reweighted l_p norm penalized LMS has a better performance than all the other algorithms. However, in the case of SNR = 10 dB, the l_0 norm constraint sparse filtering algorithm has a similar performance to the l_p norm penalized LMS, we can conclude that the l_0 norm penalized LMS may has a better performance at low SNR.



a

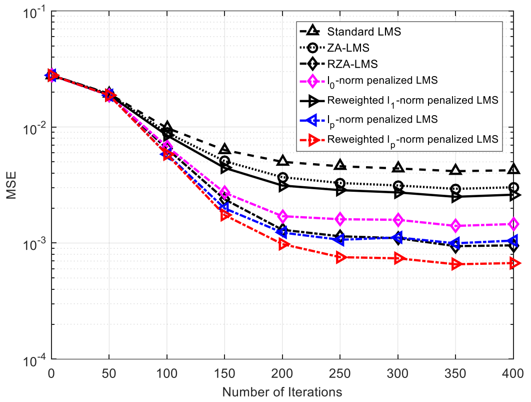


b

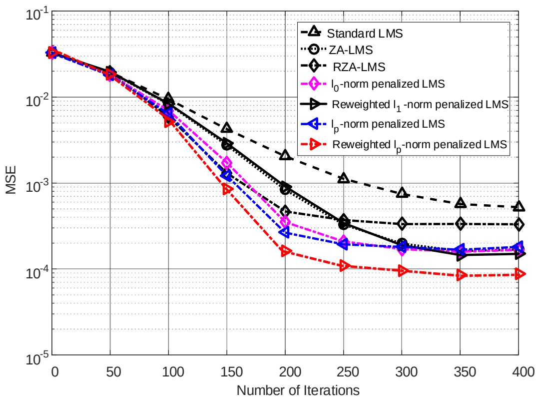
Fig. 3 Example 1, Case 2: Comparison of convergence rate for six different algorithms ($L = 31$, $d = 4$).

Experiment 2, the channel vectors have the same length $L_g = L_k = 32$ and hence the length of convolution channel vectors is $L = L_g + L_k - 1 = 63$. There are three different cases with different sparsity levels. In the first case, only two taps in $\boldsymbol{g}$ and $\boldsymbol{k}$ channel coefficients are nonzero. In the second case, four random coefficients of $\boldsymbol{g}$ and $\boldsymbol{k}$

channel impulse response are nonzero. In the third case, sixteen random channel tap-weights of the $\boldsymbol{g}$ and $\boldsymbol{k}$ are nonzero. In these three cases, all the positions of the nonzero taps in the channel coefficients vector are chosen randomly, the values of all the nonzero taps follow i.i.d Gaussian distribution. SNR of the unknown system is set to 10 dB and 20 dB, other parameters are chosen as $\mu = 0.02$, $\rho_{ZA} = \rho_{RZA} = 5 \times 10^{-4}$, $\varepsilon_{RZA} = 10$, $\rho_p = \rho_1 = 5 \times 10^{-4}$, and $\varepsilon_p = \varepsilon_1 = 1.2$.



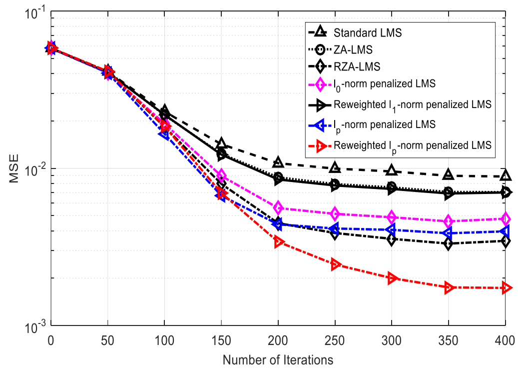
a



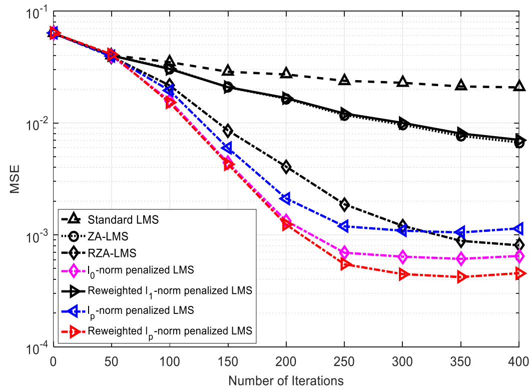
b

Fig. 4 Example 2, Case 1: Comparison of convergence rate for six different algorithms ($L = 63$,$d = 2$).
(***a***) SNR=10 dB, (**b**) SNR=20 dB

As we can see from the curves in Fig. 4 ~Fig.6, the reweighted l_p norm penalized LMS algorithm performs better and has a faster convergence rate comparing with other algorithms at low SNR. Under the same sparsity, the convergence performance in Fig.4 and Fig.5 are better than that in Fig.2 and Fig.3. It is evidence that the performance of channel estimation will be better with longer channel length under the same sparsity condition. The reason behind is that the system has a higher sparsity level in this experiment. Here we define the system sparsity level as d/L.



a

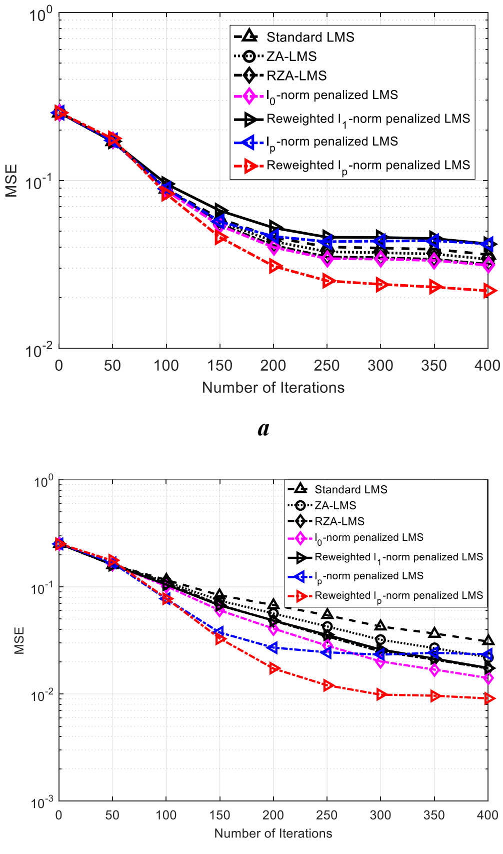


b

Fig. 5 Example 2, Case 2: Comparison of convergence rate for six different algorithms ($L = 63, d = 4$).
(***a***) SNR=10 dB, (**b**) SNR=20 dB

Experiment 3, we study the convergence of the proposed algorithm based on three different cases: various sparsity d, changing p value and different channel lengths. The simulation results are evaluated in Fig. 7~Fig. 9.

In case 1, we set the sparsity level d as 2/4/8/16 and $L_g = L_k = 32$. The positions of the nonzero taps of the channel coefficients are chosen randomly. When d has high values, the channel will have more non-zero coefficients and the system will be less sparse. In case 2, we set to p=0.4/0.5/0.7/0.9, $d = 2, L_g = L_k = 32$, for the l_p norm penalized method. In case 3, the channel length $L_g = L_k = 16/32/64$, $d = 2$.



a

b

Fig. 6 *Example 2, Case 3: Comparison of convergence rate for six different algorithms ($L = 63$, $d = 16$).* (***a***) SNR=10 dB, (**b**) SNR=20 dB

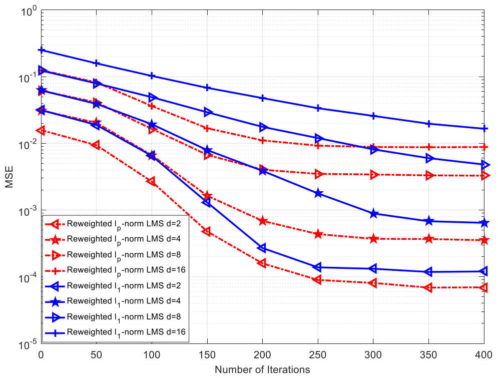


Fig. 7 tracking and convergence for two algorithms with different sparsity.

Fig. 7 shows the curves of convergence of the reweighted l_p norm penalized LMS and reweighted l_1 norm penalized LMS algorithms when the CIR sparsity level is varying. The performance of the two sparse aware LMS algorithms decreases with the increasing sparsity level of the channel, which is due to the fact that the value of $E[r_\infty]$ in (15) is increasing. The reweighted l_p norm penalized LMS algorithm achieved a better estimation performance than the reweighted l_1 norm penalized LMS algorithm. However, the performance of the reweighted l_1 norm penalized LMS algorithm has a trend to outperform the reweighted l_p norm penalized LMS algorithm at large sparsity levels.

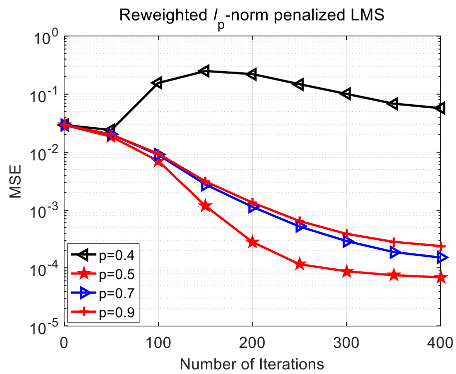


Fig. 8. *Learning curves of Reweighted l_p LMS with different p values ($d = 2, L = 63$).*

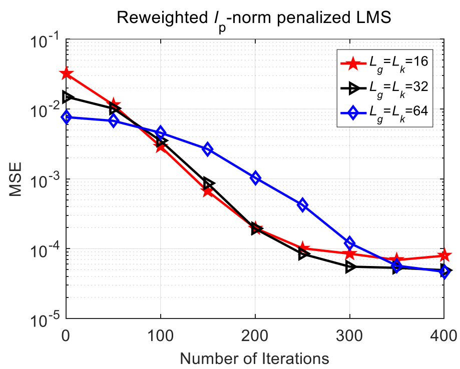


Fig. 9. Learning curves of Reweighted l_p LMS with different channel lengths. (*d*=2)

Fig. 8 shows that the estimation performance of the reweighted l_p norm penalized LMS algorithm will decrease when the p value increases, and the results indicates that the estimation has a good performance when $p = 0.5$. Fig. 9 shows that if the channel length continues to increase the estimated performance will decrease, which can be prove by the steady state bounds in (18).

5. CONCLUSIONS

A novel sparse adaptive channel estimation algorithm has been proposed in this paper for the time-variant cooperative communication systems. Cost function of the proposed method has been constructed by using reweighted l_p norm sparse penalties. Simulation results show that the proposed algorithm achieves a better convergence speed and a better steady-state behavior in comparison with other sparse aware LMS algorithms as well as the conventional LMS algorithm. We have derived the theoretical steady-state of coefficient misalignment vector and a performance upper bound. The theoretical analysis proves that the performance of the reweighted l_p norm penalized LMS algorithm is better than the performance of the reweighted l_1 norm penalized LMS algorithm.

Acknowledgment

The work was supported by the National Natural Science Foundation of China under Grant 61501530, Henan Educational Committee under Grant 16A510012, Henan science and technology planning project under Grant 182102210522.

References

[1] Laneman, J.N., Tse, D.N.C., Wornell, G.W.: 'Cooper-ative diversity in wireless networks: Efficient protocols and outage behavior', IEEE Transactions on Information Theory, 2004, 50, (12), pp. 3062-80.
[2] Zhang, A., Yang, S., Gui, G.: 'Sparse Channel Estimation for MIMO-OFDM Two-Way Relay Network with Compressed Sensing', International Journal of Antennas and Propagation, 2013, 2013, (8), pp. 72-81.
[3] Zhang, A., Yang, S., Li, J., et al.: 'Sparsity Adaptive Expectation Maximization Algorithm for Estimating Channels in MIMO Cooperation systems', KSII Transactions on Internet and Information Systems, 2016, 10, (8), pp. 3498-3511.
[4] Widrow, B., Stearns, S.D.: 'Adaptive Signal Processing', New Jersey: Prentice Hall, 1985.
[5] Jian, J., Gu, Y., Mei, S.: 'Adaptive Algorithm for Sparse System Identification: Zero-attracting LMS ', Journal of Tsinghua University, 2010, 50, (10), pp. 1656-1659.
[6] Candes, E. J., Wakin, M. B., Boyd, S. P.: 'Enhancing sparsity by reweighted $l1$ minimization ', Journal of Fourier Analysis and Applications, 2008, 14, (5), pp. 877–905.
[7] Tropp, J. A.: 'Greed is good: Algorithmic results for sparse approximation ', IEEE Trans. Information Theory, 2004, 50, (10), pp. 2231–2242.
[8] Tropp, J. A., Gilbert, A. C., 'Signal recovery from random measurements via orthogonal matching pursuit', IEEE Trans. Information Theory, 2007, 53, (12), pp. 4655–4666.
[9] Blumensath, T., Davies, M. E.: 'Normalized iterative hard thresholding: Guaranteed stability and performance', IEEE J. Selected Topics Signal Processing, 2010, 4, (2), pp. 298–309.
[10] Jin, J., Gu, Y., Mei, S.: 'New adaptive algorithm for sparse system identification: Zero-Attracting LMS', Journal of Tsinghua University Science and Technology, 2010, 50, (9), pp.1312-1315.
[11] Chen, Y., Gu, Y., Hero, A. O.: 'Sparse LMS for system identification', IEEE ICASSP, Taiwan, Apr 2009. pp. 3125-3128.
[12] Taheri, O., Vorobyov, S.A.: 'Reweighted l_1 norm penalized LMS for sparse channel estimation and its analysis', Signal Processing, 2014, 104, (6), pp.70-79.
[13] Su, G., Jin, J., Gu, Y. et al.: 'Performance Analysis of ℓ_0 Norm Constraint Least Mean Square Algorithm', IEEE Transactions on Signal Processing, 2012, 60, (5), pp.2223-2235.
[14] Hong, X., Gao, J., Chen, S., 'Zero-Attracting Recursive Least Squares Algorithms', IEEE Transactions on Vehicular Technology, 2017, 66, (1), pp. 213-221.
[15] Selesnick, I.: 'Sparse Regularization via Convex Analysis', IEEE Transactions on Signal Processing, 2017, 65, (17), pp.4481-4494.
[16] Shen, X., Gu, Y.: 'Nonconvex Sparse Logistic Regression with Weakly Convex Regularization', IEEE Transaction on Signal Processing, 2018, 66, (12), pp.3199-3211.
[17] Aliyu, M. L., Alkassim, M. A., Salman, M. S.: 'A p norm variable step-size LMS algorithm for sparse system identification ', Signal Image & Video Processing, 2015, 9, (7), pp.1559-1565.
[18] Chen, L., Gu, Y.: 'Local and Global Optimality of LP-Minimization for Sparse Recovery', IEEE International Conference on Acoustics, Speech and Signal Processing

(ICASSP), South Brisbane, August 2015, pp. 3596-3600.

Appendix

The steady state mean square derivation between the original CSI and the estimated CSI will be deduced and the condition of μ to guarantee convergence will be derived.

We define $\boldsymbol{D}(n)$ is the variance of $\boldsymbol{r}(n)$.

$$\boldsymbol{D}(n) = E[\boldsymbol{r}(n)\boldsymbol{r}^T(n)]. \quad (19)$$

Now, multiplying both sides of (10) with their respective transposes, the update equation can be expressed as

$$\begin{aligned}\boldsymbol{D}(n+1) = {} & [1 - 2\mu P_x + 2\mu^2 P_x^2] \cdot \boldsymbol{S}(n) \\ & + \mu^2 \sigma P_x^2 tr[\boldsymbol{S}(n)]\boldsymbol{I} \\ & + [1 - \mu P_x]\rho E\left[\boldsymbol{r}(n)\boldsymbol{f}(\hat{h}^T(n))\right] \\ & + [1 - \mu P_x]\rho E\left[\boldsymbol{f}(\hat{h}^T(n)\boldsymbol{r}(n))\right] \\ & + \rho^2 E\left[\boldsymbol{f}(\hat{h}(n))\boldsymbol{f}(\hat{h}^T(n))\right] + \mu^2 P_v P_x \boldsymbol{I}.\end{aligned} \quad (20)$$

Let $S(n) = tr[\boldsymbol{D}(n)]$, take the trace on both side of (20)

$$\begin{aligned}S(n+1) = {} & [1 - 2\mu P_x + (L+2)\mu^2 P_x^2] \cdot S(n) \\ & + 2[1 - \mu P_x]\rho c(n) \\ & + \rho^2 q(n) + L\mu^2 P_v P_x,\end{aligned} \quad (21)$$

where $c(n) = E\left[r^T(n)\boldsymbol{f}\left(\hat{h}(n)\right)\right], q(n) = \left\|\boldsymbol{f}(\hat{h}(n)\right\|_2^2$, $c(n)$ and $q(n)$ are all bounded and thus we can prove the condition of convergence as

$$|1 - 2\mu P_x + (\mathrm{L}+2)\mu^2 P_x| < 1.$$

Thereby we have:

$$0 < \mu < \frac{2}{(L+2)P_x}.$$

The final mean square deviation of reweighted l_p-norm penalty LMS is

$$S(\infty) = \frac{2[1-\mu P_x]\gamma c(\infty) + \gamma^2 \mu q(\infty) + L\mu P_v P_x}{P_x[2-(L+2)\mu P_x]}. \quad (21)$$

Reweighted l_p Constraint LMS-Based Adaptive Sparse Channel Estimation for Cooperative Communication System

Aihua Zhang[1*], Pengcheng Liu[2], Bing Ning[1], Qiyu Zhou[1]

1.School of Electronic and Information Engineering, Zhongyuan University of Technology, Zhengzhou, 450007, China;

2 Cardiff School of Technologies, Cardiff Metropolitan University, Cardiff CF5 2YB, UK.

*corresponding author: zhah@zut.edu.cn

Abstract: This paper studies the issue of sparsity adaptive channel reconstruction in time-varying cooperative communication networks through the amplify-and-forward transmission scheme. A new sparsity adaptive system identification method is proposed, namely reweighted l_p norm ($0 < p < 1$) penalized least mean square (LMS) algorithm. The main idea of the algorithm is to add a l_p norm penalty of sparsity into the cost function of the LMS algorithm. By doing so, the weight factor becomes a balance parameter of the associated l_p norm adaptive sparse system identification. Subsequently, the steady state of the coefficient misalignment vector is derived theoretically, with a performance upper bounds provided which serve as a sufficient condition for the LMS channel estimation of the precise reweighted l_p norm. With the upper bounds, we prove that the l_p ($0 < p < 1$) norm sparsity inducing cost function is superior to the reweighted l_1 norm. An optimal selection of p for the l_p norm problem is studied to recover various d sparse channel vectors. Several experiments verify that the simulation results agree well with the theoretical analysis, and thus demonstrate that the proposed algorithm has a better convergence speed and better steady state behavior than other LMS algorithms.

1. Introduction

Cooperative communication has been widely studied recently in wireless networks because of its significant performance in enhancing the transmission capacity and exploiting spatial diversity to against the influence of path loss and channel fading [1]. In cooperative communication systems, the accurate channel impulse response (CIR) is needed for equalization, coherent signal detection, and so on, and it can also improve the communication quality of service in 5G wireless communication systems, especially for the dynamically changing channel and its sparsity. Therefore, the issue of accurately estimate the channel state information in dynamic cooperative relay channel systems becomes significant and challenging.

In cooperative communication systems, the multipath tap-weights are spread widely in time with only a few significant components, and the impulse response of multipath wireless channel contains only a small fraction of nonzero coefficients, which means the cooperative channel has sparse structure. As such, the cooperative relay multipath wireless channel is the cooperative relay multipath wireless channel is characterized as a fast time-varying and sparse feature. By utilizing and exploiting the inherent sparsity of cooperative channel impulse response, the channel estimation performance can be improved. Currently, there has been a growing interest in sparse channel estimation, and advanced channel estimation algorithms have been developed such as compressed sensing (CS) algorithms and sparse adaptive filtering (SAF) algorithms [2–6] and so forth.

Sparse channel estimation methods mainly include: optimization methods, thresholding methods and greedy methods. Classic algorithms include basis pursuit (BP) algorithm, orthogonal matching pursuit (OMP) method and iterative thresholding algorithms [7-9]. Unfortunately, these algorithms are not applicable for sparse channel estimation in fast time-varying environments. In [3], the sparsity adaptive expectation maximization (SAEM) used expectation maximization algorithm (EM) and Kalman filter which can utilize channel sparsity well and trace the true support set of time-varying channel [3]. However, SAEM has high computational complexity.

Accordingly, the LMS-based sparse adaptive filtering or recursive least squares (RLS) algorithms are developed attribute to their simplicity in application [10–13]. In addition, a class of novel sparse adaptive algorithms has emerged based on regularized LMS algorithms, where the sparsity penalty induced strategy is used by exerting various sparsity penalty terms into the instantaneous error of a traditional adaptive filtering algorithm. The sparsity constraint can be l_1 norm [10], reweighted l_1 norm [12], l_0 norm [13], and non-convex sparsity penalty. These algorithms have good performance on faster convergence rate and smaller mean square error (MSE) comparing with the traditional adaptive filtering method, such as zero-point attraction Least Mean Square algorithm (ZA-LMS) [10], reweighted zero attracting LMS (RZA-LMS) [11] and so on. ZA-LMS uses a l_1 norm penalty in the cost function of the traditional LMS algorithm, where l_1 norm acts as a zero-point attracted term to modify the parameter vector update equation. RZA-LMS introduces the log-sum penalty and its performance is similar to the l_0 norm algorithm. Y. Gu proposed the l_0 norm Constraint LMS Algorithm [13], by exerting the l_0 norm penalty into the cost function of the LMS algorithm. The l_0 norm, a more accurate measure of sparsity, is defined as the number of non-zero elements in the unknown system vector. Similarly, as proposed in [14], ZA-RLS-I and ZA-RLS-II added l_1 norm penalty and approximated l_1 norm of the parameter vector penalty term instead of an adaptively weighted l_2 norm penalty to cost function of the RLS algorithm. The ZA-RLS algorithms achieve better performance than the other LS algorithms, however, their MSEs are not as good as sparse LMS algorithms.

Recently, the non-convex methods have received tremendous attentions in solving the problem of sparse recovery [15-18]. Furthermore, some studies have presented that the non-convex penalties might induce better sparsity than the convex penalties. In addition, the local and global optimality of l_p minimization for sparse recovery can be guaranteed even under weaker conditions in comparison with the convex l_1 minimization when the penalty approaches the l_p norm [17]. In this work, we study the fast identification of sparse cascaded channel by using the framework of an adaptive filter. In order to explore the sparse features of the cooperative relay communication system, we propose a new sparse aware LMS algorithm for relay channel reconstruction. The expectation of the misalignment vector is derived and discussed under different algorithm parameters and system sparsity. Simulation studies are conducted to verify the high robustness, low computational cost and easy implementation of the proposed algorithm.

This paper is organized as follows. The amplify-and-forward (AF)-based cooperative relay channel model is described briefly in Section 2. Then in Section 3, we introduce the reweighted l_p norm constraint LMS and derive the expectation of the misalignment vector and provide the steady-state analysis of the proposed algorithm. Numerical simulations and rigorous analysis are presented in Section 4 to demonstrate the effectiveness prove the theoretical analysis. Finally, the conclusion is given in section 5.

2. System model and LMS algorithm

2.1 Cooperative Rely Channel Model

Consider an amplified model of cooperation relay network with a source node $\mathbf{T_1}$, a destination node $\mathbf{T_2}$, and one relay node $\mathbf{R}$. It is assumed that all the terminals are equipped with only one antenna and work in the half-duplex mode. When node $\mathbf{T_2}$ is beyond the communication range of node $\mathbf{T_1}$ duo to remote distance or shielding affection, then all signals sent by the source $\mathbf{T_1}$ need to be forwarded to destination $\mathbf{T_2}$ by relay node. Denote $\boldsymbol{g} = [g_0, g_1, \cdots g_{L_g-1}]$ as the baseband channel between $\mathbf{T_1}$ and $\mathbf{R}$. And $\boldsymbol{k} = [k_0, k_1, \cdots k_{L_k-1}]$ is the channel vector between relay node $\mathbf{R}$ and destination $\mathbf{T_2}$. Since $\mathbf{T_1}$ and $\mathbf{T_2}$ are separated from each other, $\boldsymbol{g}$ and $\boldsymbol{k}$ are considered independent. The taps of all these two channels are assumed as zero-mean circularly symmetric complex Gaussian random variables, i.e., $g_i \sim \mathcal{CN}(0, \sigma_{g,i}^2)$, $k_i \sim \mathcal{CN}(0, \sigma_{k,i}^2)$. Moreover, the source and relay are assumed to have average power constraints, which are denoted by P_1 and P_R, respectively.

There are two stages in amplified relay transmission system, and it takes two time slots to achieve the cooperative multiple access. In the first time slot, the source node $\mathbf{T_1}$ sends signals $\boldsymbol{x}$ and the relay node $\mathbf{R}$ receives as

$$\boldsymbol{r} = \boldsymbol{g}\boldsymbol{x} + \boldsymbol{n}_r, \tag{1}$$

where $\boldsymbol{n}_r$ is the additive white Gaussian noise with variance σ_r^2.

In the second time slot, the relay node amplifies and transmits the received data to the destination node $\mathbf{T_2}$, and $\mathbf{T_2}$ receives as

$$\boldsymbol{y}(n) = \underbrace{\alpha \boldsymbol{k}\boldsymbol{g}}_{\boldsymbol{h}}\, \boldsymbol{x} + \underbrace{\alpha \boldsymbol{k}\boldsymbol{n}_r + \boldsymbol{n}_1}_{\boldsymbol{n}}$$

$$= \boldsymbol{hx} + \boldsymbol{n}, \quad (2)$$

where $\boldsymbol{h}$ (with the length of $L = L_g + L_k - 1$) is the cascaded channel that is the convolution between $\boldsymbol{g}$ and $\boldsymbol{k}$, $\boldsymbol{n}_1$ represents the noise at $\mathbf{T}_2$ with variance σ_1^2, $\boldsymbol{n}$ denotes the overall noise. $\alpha = \sqrt{P_R/[P_1\sigma_g^2 + \sigma_r^2]}$ and $\sigma_g^2 = \sum_{i=0}^{L_g-1}\sigma_{g,i}^2$.

2.2 Standard LMS

In the AF relay cooperative communication system, the unknown cascaded channel coefficients at time instant n are $\boldsymbol{h} = [h_0, h_1, \cdots, h_{L-1}]^T$. The system's input data vector from $\mathbf{T}_1$ is expressed as $\boldsymbol{x} = [x_n, x_{n-1}, \cdots, x_{n-L+1}]^T$ and it is assumed to be independent Gaussian input. As shown in Fig. 1, we consider sparse adaptive channel estimation in a relay-based cooperation communication system.

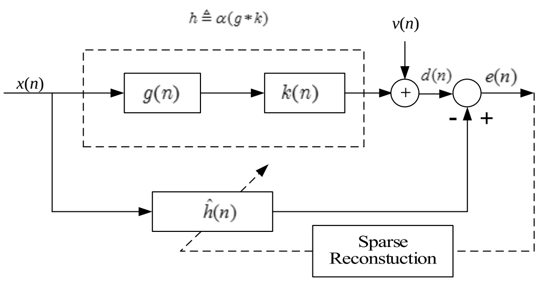


Fig. 1 Channel model and adaptive channel estimation algorithm

In Fig. 1, the desired output signal $\boldsymbol{d}(n)$ accord to $\boldsymbol{y}(n)$, denoted as

$$\boldsymbol{d}(n) = \boldsymbol{h}^T(n)\boldsymbol{x}(n) + \boldsymbol{v}(n), \quad (3)$$

where $\boldsymbol{v}(n)$ is the noise signal. The estimated error $\boldsymbol{e}(n)$ is the instantaneous error between the output signal of the unknown system and the output from the adaptive filter, which can be written as

$$\boldsymbol{e}(n) = \boldsymbol{d}(n) - \widehat{\boldsymbol{h}}^T(n)\boldsymbol{x}(n), \quad (4)$$

where $\widehat{\boldsymbol{h}} = [\widehat{\boldsymbol{h}}_0, \widehat{\boldsymbol{h}}_1, \cdots \widehat{\boldsymbol{h}}_{L-1}]^T$ is defined as an adaptive tap-weights vector. The noise vector $\boldsymbol{v}(n)$ follows i.i.d. zero mean and δ_n^2 variance white Gaussian distribution. It is assumed that the adaptive tap-weights vector $\boldsymbol{h}(n)$, input signal $\boldsymbol{x}(n)$ and additive noise signal $\boldsymbol{v}(n)$ are mutually independent.

According to the standard LMS framework, the cost function is defined as $\xi(n) = 0.5|\boldsymbol{e}(n)|^2$. The recursive equation of the filter coefficient vector can be derived as

$$\widehat{\boldsymbol{h}}(n+1) = \widehat{\boldsymbol{h}}(n) - \mu\frac{\partial\xi(n)}{\partial\widehat{\boldsymbol{h}}(n)} = \widehat{\boldsymbol{h}}(n) + \mu\boldsymbol{e}(n)\boldsymbol{x}(n), \quad (5)$$

where μ is the step size parameter which satisfies $\mu \in (0, 1/\lambda_{max})$ and adjusts the convergence rate and the steady performance behavior of the LMS algorithm[4]. λ_{max} is the maximum eigenvalue of P_x, here $P_x = E[\boldsymbol{x}(n)\boldsymbol{x}^T(n)]$, which denotes the covariance matrix of the input vector $\boldsymbol{x}(n)$.

3. The proposed channel estimation algorithm

This work focuses on fast unknown channel identification of the cooperation system via sparse constraint adaptive filter. The impulse response of a sparse cooperative system consists of few nonzero coefficients, most of the coefficients in the channel representation vector of $\boldsymbol{h}(n)$ in time delay domain should be zeroes or small values. In order to improve the performance of sparsity adaptive channel estimation, we propose a novel sparsity-aware system identification method with a new cost function.

3.1 Reweighted l_p Norm Penalized LMS

For the sake of exploiting the sparse structure of cooperative relay communication system, we propose an idea of introducing reweighted l_p norm of channel impulse response as a sparsity penalty into the cost function of traditional LMS algorithm. This method can accelerate and enhance the performance of sparse cascade channel estimation.

Motivated by the research findings that non-convex penalties might induce better sparsity than the convex penalties [18], we apply reweighted l_p $(0 < p < 1)$ norm to measure the sparsity of channel vector.

$$\left\|\widehat{\boldsymbol{h}}(n)\right\|_p = \left(\sum_i\left|\widehat{\boldsymbol{h}}(n)\right|^p\right)^{\frac{1}{p}}, \quad (6)$$

where $\|\cdot\|_p$ stands for the l_p norm of the channel vector.

When $0 < p < 1$

$$\lim_{p\to1}\left\|\widehat{\boldsymbol{h}}(n)\right\|_p = \left\|\widehat{\boldsymbol{h}}(n)\right\|_1 = \sum_{i=1}^{L}\left|\widehat{\boldsymbol{h}}(i)\right|.$$

$$\lim_{p\to0}\left\|\widehat{\boldsymbol{h}}(n)\right\|_p = \left\|\widehat{\boldsymbol{h}}(n)\right\|_0.$$

In order to learn the time-varying channel state information by using the prior sparsity, we apply a new cost function, which combines the instantaneous channel estimation square error and the l_p $(0 < p < 1)$ norm penalty term of the coefficient vector, defined as

$$\xi_p(n) = \frac{1}{2}|\boldsymbol{e}(n)|^2 + \gamma_p\left\|\widehat{\boldsymbol{h}}(n)\right\|_p, \quad (7)$$

where γ_p can be selected using a positive factor to balance the mean square error and adjust the penalty of l_p norm.

The gradient of $\xi_p(n)$ is

$$\frac{\partial \xi_p(n)}{\partial \hat{\boldsymbol{h}}(n)} = -\boldsymbol{e}(n)\boldsymbol{x}(n) + \gamma_p \frac{\left(\|\hat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)} sgn(\hat{\boldsymbol{h}}(n))}{|\hat{\boldsymbol{h}}(n)|^{(1-p)}}. \quad (8)$$

On the right-hand side of equation (8), the second term is the gradient of l_p norm.

Following the gradient descent algorithm, we use the gradient of l_p norm as the zero attracting term. The reweighted l_p norm LMS algorithm updates its coefficients by (9)

$$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu \boldsymbol{e}(n)\boldsymbol{x}(n) - \rho_p \frac{\left(\|\hat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)} \mathrm{sgn}(\hat{\boldsymbol{h}}(n))}{\varepsilon_p + |\hat{\boldsymbol{h}}(n)|^{(1-p)}}, \quad (9)$$

where $\rho_p = \alpha\mu\gamma_p$ with $\alpha\ (0 < \alpha < 1)$ is a reweighted parameter adopted to adjust zero-point attraction ability of the l_p norm penalty. We set parameter $\varepsilon_p > 0$ for providing stability and ensuring that a zero value in $\hat{h}(n)$ does not prohibit a non-zero estimate strictly [6]. Additionally, ε_p should be set as a small positive value or smaller than the expected nonzero magnitudes of $\hat{h}(n)$. In consequence, the reasonably robust channel estimation process tends to be dependent on the choice of ε_p.

The procedure of the reweighted l_p norm LMS algorithm is listed in Table 1, where 0_L is a zero vector of size L.

Table 1 The proposed reweighted l_p norm LMS algorithm

Require: $\mu > 0, \gamma > 0, \alpha > 0,\ \varepsilon_p > 0\ ,\ 0 < p < 1,$
$\boldsymbol{g}$ and $\boldsymbol{k}$ as zeroes or small random vectors;

1. **Initialize** $\hat{h}(0) = 0_L, e(0) = 0,\ i=1$;
2. while $i < N$
3. Calculate the error through equation (4);
4. Update the gradient of $\xi_p(n)$ according to equation (7);
5. Update the reweighted zero attractor by multiplying α

$$\alpha \times \mu\gamma_p \times \left(\|\hat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)} \times \mathrm{sgn}(\hat{\boldsymbol{h}}(n)/[\varepsilon_p + |\hat{\boldsymbol{h}}(n)|^{(1-p)}].$$

6. Update the tap-weight vector according to equation (9)
7. $i = i + 1$;
8. end while

Reweighted l_1 norm penalty: Another method of exploring the sparsity of the wireless communication channel is to use a reweighted l_1 norm constraint in the cost function of SAF algorithms [12]. The reweighted l_1 norm needs to be minimized, which approximates the l_0 pseudo norm better than the l_1 norm. The reweighted l_1 norm penalized LMS algorithm considers a l_1 norm penalty term to channel impulse response vector. The reweighted l_1 norm cost function in [12] is derived as

$$\xi_1(n) = \frac{1}{2}|\boldsymbol{e}(n)|^2 + \gamma_1 \left\|\boldsymbol{s}(n)\hat{\boldsymbol{h}}(n)\right\|_1,$$

where γ_1 is the weight for adjusting the l_1 norm penalty term. And $\boldsymbol{s}(n)$ acts as a weight element that can be denoted as

$$[\boldsymbol{s}(n)]_i = \frac{1}{\varepsilon_1 + \left|[\hat{\boldsymbol{h}}(n-1)]_i\right|}, i = 1,2,\cdots,L.$$

where ε_1 should be set as a small positive value and $[\cdot]_i$ is the ith entry of the estimated channel coefficient vector. The gradient of $\xi_1(n)$ can be written as

$$\frac{\partial \xi_1(n)}{\partial \hat{\boldsymbol{h}}(n)} = -\boldsymbol{e}(n)\boldsymbol{x}(n) + \gamma_1 \frac{\mathrm{sgn}(\hat{\boldsymbol{h}}(n))}{\varepsilon_1 + |\hat{\boldsymbol{h}}(n-1)|}.$$

The resulting update is

$$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu \boldsymbol{e}(n)\boldsymbol{x}(n) - \rho_1 \frac{\mathrm{sgn}(\hat{\boldsymbol{h}}(n))}{\varepsilon_1 + |\hat{\boldsymbol{h}}(n-1)|},$$

where $\rho_1 = \mu\gamma_1$.

3.2. Computational Complexity

Computational complexity of the proposed reweighted l_p Constraint LMS algorithm and some sparsity-aware LMS algorithms are compared in Table 2, in terms of arithmetic operations and comparisons. As shown in Table 2, the proposed algorithm has lower computation complexity than l_0-LMS [13]. And the amount of computations of the proposed algorithm is similar to that of RZA-LMS and the reweighted l_1 norm penalized LMS.

Table 2 Computational complex of different algorithms

Algorithm	Computational Complexity
LMS	(2L) Add+(2L+1) Multiply
RZA-LMS	(4L) Add+(5L+1) Multiply
l_0-LMS	(4L) Add+(5L+1) Multiply+(L) Comp
Reweighted l_1-LMS	(4L) Add+(5L+1) Multiply
Proposed algorithm	(4L) Add+(5L+1) Multiply

3.3 Performance Analysis

In this section, we derive the theoretical steady-state of the coefficient misalignment and provide a mean square error convergence analysis of the new sparsity adaptive channel estimation algorithm. Then a performance upper bounds is drawn as a sufficient condition for the precise reweighted l_p norm LMS channel estimation.

A. Mean Performance

Assuming an i.i.d. zero-mean Gaussian input signal $\boldsymbol{x}(n)$ and a zero mean white noise. We define $\boldsymbol{r}(n) = \hat{\boldsymbol{h}}(n) - \boldsymbol{h}(n)$ as the filter misalignment vector. The recursion formula of the misalignment vector can be written as

$$\boldsymbol{r}_{n+1} = (\boldsymbol{I} - \mu \boldsymbol{x}_n \boldsymbol{x}_n^T) r_n + \mu \boldsymbol{v}_n \boldsymbol{x}_n - \rho_p f(\hat{\boldsymbol{h}}_n), \quad (10)$$

where $f(\hat{\boldsymbol{h}}_n)$ is defined as

$$f(\hat{\boldsymbol{h}}_n) = \frac{\left(\|\hat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)} \operatorname{sgn}(\hat{\boldsymbol{h}}(n))}{\varepsilon_p + |\hat{\boldsymbol{h}}(n)|^{(1-p)}}. \quad (11)$$

Taking expectation, since $\boldsymbol{v}_n$ is assumed to have a zero mean and be independent with input signal $\boldsymbol{x}(n)$. Assume that $\boldsymbol{h}(n)$ is a d-sparse channel vector, we have

$$E[\boldsymbol{r}_{n+1}] = (I - \mu P_x) E[\boldsymbol{r}_n] - \rho_p E[f(\hat{\boldsymbol{h}}_n)]. \quad (12)$$

$$E[f(\hat{\boldsymbol{h}}_n)] = E[\frac{\left(\|\hat{\boldsymbol{h}}(n)\|_p\right)^{(1-p)} \operatorname{sgn}(\hat{\boldsymbol{h}}(n))}{\varepsilon_p + |\hat{\boldsymbol{h}}(n)|^{(1-p)}}]. \quad (13)$$

$$-\frac{\rho_p}{\mu P_x}\mathbf{1} < E[r_{\infty}] < \frac{\rho_p}{\mu P_x}\mathbf{1}. \quad (14)$$

Further derivation, we obtain

$$-E[r_{\infty}] < \frac{\rho_p}{\mu P_x}\mathbf{1} \le \frac{\rho_p(\sqrt[p]{d})^{1-p}}{\mu P_x[\varepsilon_p + |\hat{\boldsymbol{h}}(n)|^{(1-p)}]}\mathbf{1}, \quad (15)$$

where d denotes the number of non-zero coefficients, which means the sparsity level of cooperative communication channel. It can be observed that $E[r_{\infty}]$ is bounded between $-\rho_p \mathbf{1}/(\mu P_x)$ and $\rho_p \mathbf{1}/(\mu P_x)$, where $\mathbf{1}$ is the vector with all one entries. This means that the reweighted l_p norm LMS algorithm has a stability condition for the coefficient misalignment vector convergence. From (15), we can achieve better performance by adjusting parameter ε_p by following the change of sparsity d.

The mean misalignment vector of the reweighted l_1 norm penalized LMS algorithm is bounded as

$$-\frac{\rho_1}{\mu P_x \varepsilon_r}\mathbf{1} \le E[r_{\infty}] \le \frac{\rho_1}{\mu P_x \varepsilon_r}\mathbf{1}. \quad (16)$$

when $\rho_1 = \rho_p$, as ε_r is a very small positive value, so $\rho_p \mathbf{1}/(\mu P_x) < \rho_1 \mathbf{1}/(\mu P_x \varepsilon_r)$ generally. Theoretically, the performance of reweighted l_p norm sparse aware LMS algorithm is better than the reweighted l_1 norm penalized LMS algorithm.

B. Mean Square Steady-State Performance of proposed algorithm

The mean square deviation (MSD) bounds of the proposed reweighted l_p norm constraint LMS are derived by the following theorem. In order to guarantee convergence, step-size μ should satisfy

$$0 < \mu < \frac{2}{(L+2)P_x}. \quad (17)$$

The final mean square deviation of the proposed algorithm is

$$S(\infty) = \frac{2[1-\mu P_x]\gamma c(\infty) + \gamma^2 \mu q(\infty) + L\mu P_v P_x}{P_x[2-(L+2)\mu P_x]}, \quad (18)$$

where $c(n) = E\left[r^T(n)\boldsymbol{f}\left(\hat{h}(n)\right)\right]$, $q(n) = \left\|\boldsymbol{f}(\hat{h}(n)\right\|_2^2$, $c(n)$ and $q(n)$ are all bounded. The proof of (17) and (18) goes in Appendix.

From (17), we can conclude that the MSD will decrease as the channel length L increases, which has been proved in the simulations. It is shown that the convergence of the proposed algorithm can be guaranteed when μ satisfies (17). When p closes to zero, steady-state MSD of our proposed algorithm will be smaller and then the steady state performance will be better than other sparse LMS algorithm.

4. Simulation results and analysis

In this section, three experiments are provided to demonstrate the estimation performance of sparse adaptive filtering algorithms. The reference algorithms simulated for comparison with the proposed algorithm include Standard LMS [4], ZA-LMS [10], RZA-LMS [11], reweighted l_1 norm penalized LMS algorithms [12], and l_0 LMS [13]. The cost functions and updated equations of the above algorithms are listed in Table 3.

Table 3 Various LMS algorithms

Algorithm	Cost function	Update equation
LMS	$\xi(n) = \frac{1}{2}\lvert\boldsymbol{e}(n)\rvert^2$	$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu\boldsymbol{e}(n)\boldsymbol{x}(n)$
ZA-LMS	$\xi_{ZA}(n) = \frac{1}{2}\lvert\boldsymbol{e}(n)\rvert^2 + \gamma_{ZA}\lVert\hat{\boldsymbol{h}}(n)\rVert$	$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu\boldsymbol{e}(n)\boldsymbol{x}(n)$ $-\rho_{ZA}\operatorname{sgn}[\hat{\boldsymbol{h}}(n)]$
RZA-LMS	$\xi_{RZA}(n) = \frac{1}{2}\lvert\boldsymbol{e}(n)\rvert^2$ $+\gamma_{RZA}\sum_{l=1}^{L}\log(1+\hat{\boldsymbol{h}}(n))$	$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu\boldsymbol{e}(n)\boldsymbol{x}(n)$ $-\rho_{RZA}\frac{\operatorname{sgn}[\hat{\boldsymbol{h}}(n)]}{1+\varepsilon_{RZA}\lvert\hat{\boldsymbol{h}}(n)\rvert}$
l_0-LMS	$\xi_0(n) = \frac{1}{2}\lvert\boldsymbol{e}(n)\rvert^2 + \gamma_0\lVert\hat{\boldsymbol{h}}(n)\rVert_0$	$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu\boldsymbol{e}(n)\boldsymbol{x}(n)$ $-\rho_0\beta\operatorname{sgn}[\hat{\boldsymbol{h}}(n)]e^{-\beta\lvert\hat{\boldsymbol{h}}(n)\rvert}$
Proposed algorithm	$\xi_p(n) = \frac{1}{2}\lvert\boldsymbol{e}(n)\rvert^2 + \gamma_p\lVert\hat{\boldsymbol{h}}(n)\rVert_p$	$\hat{\boldsymbol{h}}(n+1) = \hat{\boldsymbol{h}}(n) + \mu\boldsymbol{e}(n)\boldsymbol{x}(n)$ $-\rho_p\frac{\left(\lVert\hat{\boldsymbol{h}}(n)\rVert_p\right)^{(1-p)}\operatorname{sgn}(\hat{\boldsymbol{h}}(n))}{\varepsilon_p+\lvert\hat{\boldsymbol{h}}(n)\rvert^{(1-p)}}$

Experiment 1, we assume that the channel vectors of cooperative relay system have the same length $L_g = L_k = 16$ and hence the length of convolution channel vectors is $L = L_g + L_k - 1 = 31$. In case 1, two large coefficients of $\boldsymbol{g}$ and $\boldsymbol{k}$ are uniformly distributed and all the others are exactly zero, making the system have a sparsity level of $^2/_{31}$. Four random tap-weights of $\boldsymbol{g}$ and $\boldsymbol{k}$ are nonzero in case 2.

The values of several large coefficients are chosen from Gaussian distribution with a zero mean and a unit variance.

The parameters of reweighted l_p norm LMS channel estimation algorithm and reweighted l_1 norm LMS algorithm are set to $\rho_p = \rho_1 = 5 \times 10^{-4}$ and $\varepsilon_p = \varepsilon_1 = 1.2$. In the ZA-LMS algorithm and RZA-LMS algorithm, the parameters are set as $\rho_{ZA} = \rho_{RZA} = 5 \times 10^{-4}$ and $\varepsilon_{RZA} = 10$. We set step size parameter $\mu = 0.02$ for all algorithms in this paper. We test the performance of the proposed algorithm under the low Signal-to-Noise Ratio (SNR) 10dB and high SNR 20dB respectively. The average estimation of the mean square error (MSE) between the actual and estimated channel state information are shown in Fig. 2 and Fig. 3.

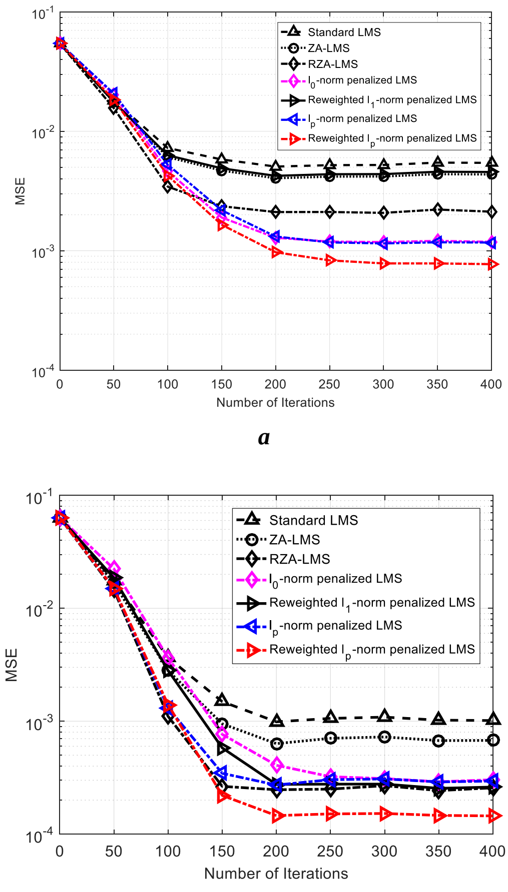


Fig. 2 Example 1, Case 1: Comparison of convergence rate for six different algorithms ($L = 31$, $d = 2$).
(***a***) SNR=10 dB, (**b**) SNR=20 dB

The estimated channel impulse response MSE results for sparsity case 1 are shown in Fig. 2 (a)~ (b) and for sparsity case 2 are shown in Fig. 3 (a) ~ (b). According to Figs. 2~3, when the sparsity of the channel increases, the convergence performance of the sparsity-aware parameter estimation algorithms degrades accordingly. By examining the convergence lines, we can conclude that, in general, the reweighted l_p norm penalized LMS has a better performance than all the other algorithms. However, in the case of SNR = 10 dB, the l_0 norm constraint sparse filtering algorithm has a similar performance to the l_p norm penalized LMS, we can conclude that the l_0 norm penalized LMS may has a better performance at low SNR.

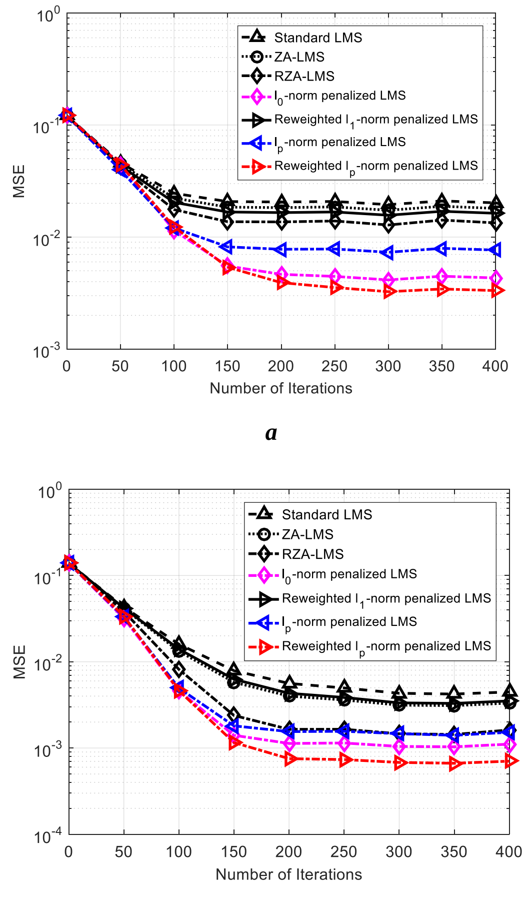
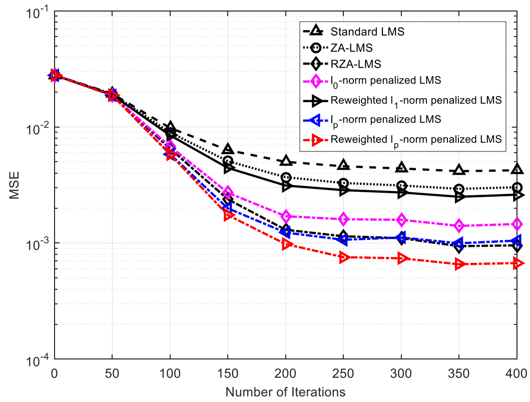


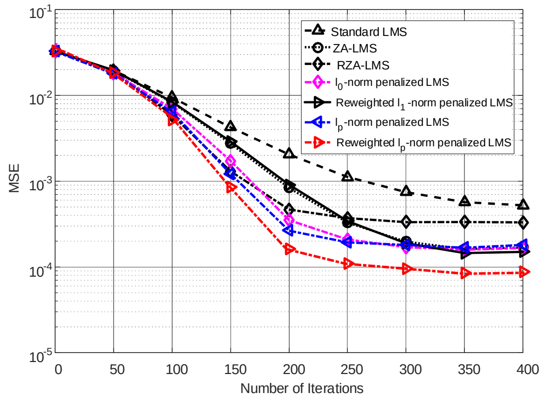
Fig. 3 Example 1, Case 2: Comparison of convergence rate for six different algorithms ($L = 31$, $d = 4$).

Experiment 2, the channel vectors have the same length $L_g = L_k = 32$ and hence the length of convolution channel vectors is $L = L_g + L_k - 1 = 63$. There are three different cases with different sparsity levels. In the first case, only two taps in $\boldsymbol{g}$ and $\boldsymbol{k}$ channel coefficients are nonzero. In the second case, four random coefficients of $\boldsymbol{g}$ and $\boldsymbol{k}$

channel impulse response are nonzero. In the third case, sixteen random channel tap-weights of the $\boldsymbol{g}$ and $\boldsymbol{k}$ are nonzero. In these three cases, all the positions of the nonzero taps in the channel coefficients vector are chosen randomly, the values of all the nonzero taps follow i.i.d Gaussian distribution. SNR of the unknown system is set to 10 dB and 20 dB, other parameters are chosen as $\mu = 0.02$, $\rho_{ZA} = \rho_{RZA} = 5 \times 10^{-4}$, $\varepsilon_{RZA} = 10$, $\rho_p = \rho_1 = 5 \times 10^{-4}$, and $\varepsilon_p = \varepsilon_1 = 1.2$.



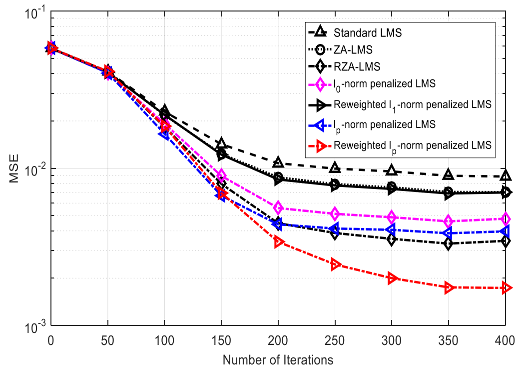
a



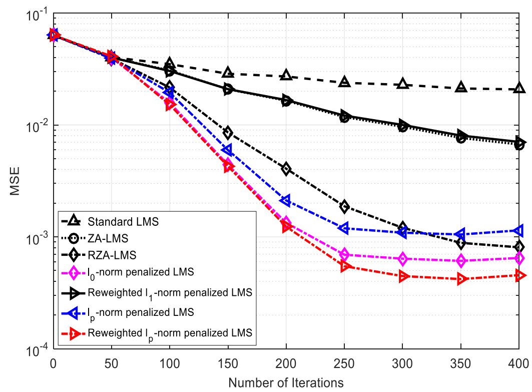
b

Fig. 4 Example 2, Case 1: Comparison of convergence rate for six different algorithms ($L = 63$, $d = 2$).
(***a***) SNR=10 dB, (**b**) SNR=20 dB

As we can see from the curves in Fig. 4 ~Fig.6, the reweighted l_p norm penalized LMS algorithm performs better and has a faster convergence rate comparing with other algorithms at low SNR. Under the same sparsity, the convergence performance in Fig.4 and Fig.5 are better than that in Fig.2 and Fig.3. It is evidence that the performance of channel estimation will be better with longer channel length under the same sparsity condition. The reason behind is that the system has a higher sparsity level in this experiment. Here we define the system sparsity level as d/L.



a

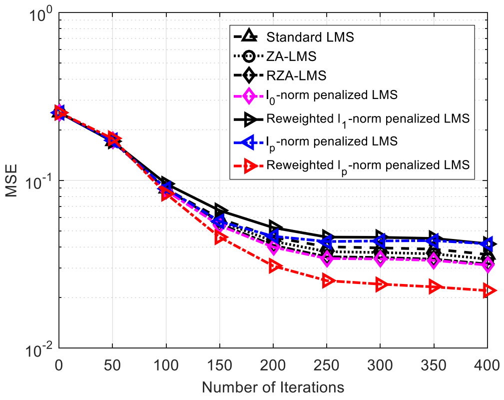


b

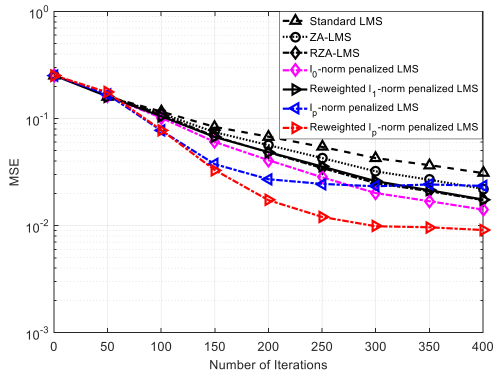
Fig. 5 Example 2, Case 2: Comparison of convergence rate for six different algorithms ($L = 63$, $d = 4$).
(***a***) SNR=10 dB, (**b**) SNR=20 dB

Experiment 3, we study the convergence of the proposed algorithm based on three different cases: various sparsity d, changing p value and different channel lengths. The simulation results are evaluated in Fig. 7~Fig. 9.

In case 1, we set the sparsity level d as 2/4/8/16 and $L_g = L_k = 32$. The positions of the nonzero taps of the channel coefficients are chosen randomly. When d has high values, the channel will have more non-zero coefficients and the system will be less sparse. In case 2, we set to p=0.4/0.5/0.7/0.9, $d = 2, L_g = L_k = 32$, for the l_p norm penalized method. In case 3, the channel length $L_g = L_k = 16/32/64$, $d = 2$.



a



b

Fig. 6 Example 2, Case 3: Comparison of convergence rate for six different algorithms ($L = 63$, $d = 16$). (***a***) SNR=10 dB, (***b***) SNR=20 dB

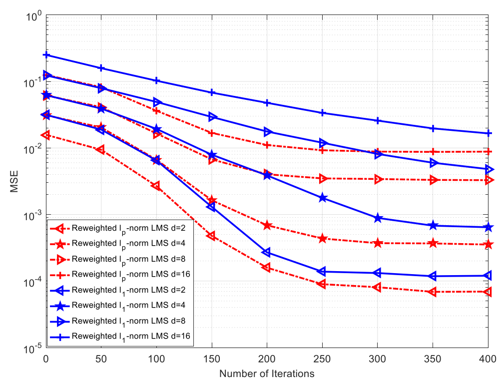


Fig. 7 tracking and convergence for two algorithms with different sparsity.

Fig. 7 shows the curves of convergence of the reweighted l_p norm penalized LMS and reweighted l_1 norm penalized LMS algorithms when the CIR sparsity level is varying. The performance of the two sparse aware LMS algorithms decreases with the increasing sparsity level of the channel, which is due to the fact that the value of $E[r_\infty]$ in (15) is increasing. The reweighted l_p norm penalized LMS algorithm achieved a better estimation performance than the reweighted l_1 norm penalized LMS algorithm. However, the performance of the reweighted l_1 norm penalized LMS algorithm has a trend to outperform the reweighted l_p norm penalized LMS algorithm at large sparsity levels.

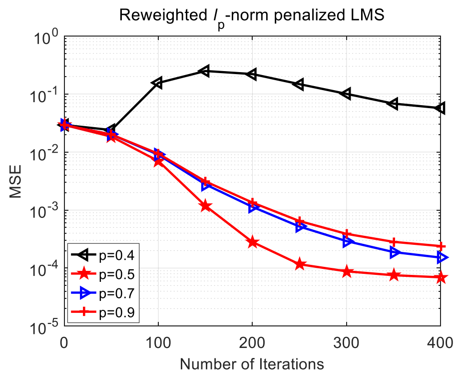


Fig. 8. Learning curves of Reweighted l_p LMS with different p values ($d = 2, L = 63$).

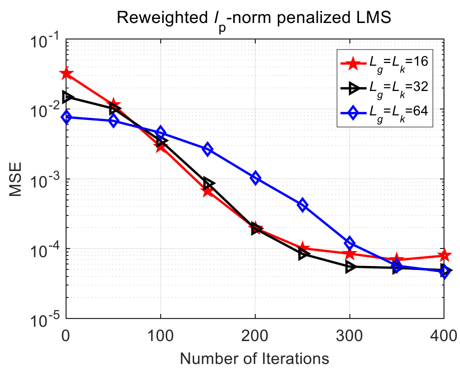


Fig. 9. Learning curves of Reweighted l_p LMS with different channel lengths. (d=2)

Fig. 8 shows that the estimation performance of the reweighted l_p norm penalized LMS algorithm will decrease when the p value increases, and the results indicates that the estimation has a good performance when $p = 0.5$. Fig. 9 shows that if the channel length continues to increase the estimated performance will decrease, which can be prove by the steady state bounds in (18).

5. CONCLUSIONS

A novel sparse adaptive channel estimation algorithm has been proposed in this paper for the time-variant cooperative communication systems. Cost function of the proposed method has been constructed by using reweighted l_p norm sparse penalties. Simulation results show that the proposed algorithm achieves a better convergence speed and a better steady-state behavior in comparison with other sparse aware LMS algorithms as well as the conventional LMS algorithm. We have derived the theoretical steady-state of coefficient misalignment vector and a performance upper bound. The theoretical analysis proves that the performance of the reweighted l_p norm penalized LMS algorithm is better than the performance of the reweighted l_1 norm penalized LMS algorithm.

Acknowledgment

The work was supported by the National Natural Science Foundation of China under Grant 61501530, Henan Educational Committee under Grant 16A510012, Henan science and technology planning project under Grant 182102210522.

References

[1] Laneman, J.N., Tse, D.N.C., Wornell, G.W.: 'Cooperative diversity in wireless networks: Efficient protocols and outage behavior', IEEE Transactions on Information Theory, 2004, 50, (12), pp. 3062-80.

[2] Zhang, A., Yang, S., Gui, G.: 'Sparse Channel Estimation for MIMO-OFDM Two-Way Relay Network with Compressed Sensing', International Journal of Antennas and Propagation, 2013, 2013, (8), pp. 72-81.

[3] Zhang, A., Yang, S., Li, J., et al.: 'Sparsity Adaptive Expectation Maximization Algorithm for Estimating Channels in MIMO Cooperation systems', KSII Transactions on Internet and Information Systems, 2016, 10, (8), pp. 3498-3511.

[4] Widrow, B., Stearns, S.D.: 'Adaptive Signal Processing', New Jersey: Prentice Hall, 1985.

[5] Jian, J., Gu, Y., Mei, S.: 'Adaptive Algorithm for Sparse System Identification: Zero-attracting LMS ', Journal of Tsinghua University, 2010, 50, (10), pp. 1656-1659.

[6] Candes, E. J., Wakin, M. B., Boyd, S. P.: 'Enhancing sparsity by reweighted $l1$ minimization ', Journal of Fourier Analysis and Applications, 2008, 14, (5), pp. 877–905.

[7] Tropp, J. A.: 'Greed is good: Algorithmic results for sparse approximation ', IEEE Trans. Information Theory, 2004, 50, (10), pp. 2231–2242.

[8] Tropp, J. A., Gilbert, A. C., 'Signal recovery from random measurements via orthogonal matching pursuit', IEEE Trans. Information Theory, 2007, 53, (12), pp. 4655–4666.

[9] Blumensath, T., Davies, M. E.: 'Normalized iterative hard thresholding: Guaranteed stability and performance', IEEE J. Selected Topics Signal Processing, 2010, 4, (2), pp. 298–309.

[10] Jin, J., Gu, Y., Mei, S.: 'New adaptive algorithm for sparse system identification: Zero-Attracting LMS', Journal of Tsinghua University Science and Technology, 2010, 50, (9), pp.1312-1315.

[11] Chen, Y., Gu, Y., Hero, A. O.: 'Sparse LMS for system identification', IEEE ICASSP, Taiwan, Apr 2009. pp. 3125-3128.

[12] Taheri, O., Vorobyov, S.A.: 'Reweighted l_1 norm penalized LMS for sparse channel estimation and its analysis', Signal Processing, 2014, 104, (6), pp.70-79.

[13] Su, G., Jin, J., Gu, Y. et al.: 'Performance Analysis of l_0 Norm Constraint Least Mean Square Algorithm', IEEE Transactions on Signal Processing, 2012, 60, (5), pp.2223-2235.

[14] Hong, X., Gao, J., Chen, S., 'Zero-Attracting Recursive Least Squares Algorithms', IEEE Transactions on Vehicular Technology, 2017, 66, (1), pp. 213-221.

[15] Selesnick, I.: 'Sparse Regularization via Convex Analysis', IEEE Transactions on Signal Processing, 2017, 65, (17), pp.4481-4494.

[16] Shen, X., Gu, Y.: 'Nonconvex Sparse Logistic Regression with Weakly Convex Regularization', IEEE Transaction on Signal Processing, 2018, 66, (12), pp.3199-3211.

[17] Aliyu, M. L., Alkassim, M. A., Salman, M. S.: 'A p norm variable step-size LMS algorithm for sparse system identification ', Signal Image & Video Processing, 2015, 9, (7), pp.1559-1565.

[18] Chen, L., Gu, Y.: 'Local and Global Optimality of LP-Minimization for Sparse Recovery', IEEE International Conference on Acoustics, Speech and Signal Processing

(ICASSP), South Brisbane, August 2015, pp. 3596-3600.

Appendix

The steady state mean square derivation between the original CSI and the estimated CSI will be deduced and the condition of μ to guarantee convergence will be derived.

We define $\boldsymbol{D}(n)$ is the variance of $\boldsymbol{r}(n)$.

$$\boldsymbol{D}(n) = E[\boldsymbol{r}(n)\boldsymbol{r}^T(n)]. \tag{19}$$

Now, multiplying both sides of (10) with their respective transposes, the update equation can be expressed as

$$\begin{aligned}\boldsymbol{D}(n+1) = {} & [1 - 2\mu P_x + 2\mu^2 P_x^2] \cdot \boldsymbol{S}(n) \\ & +\mu^2 \sigma P_x^2 tr[\boldsymbol{S}(n)]\boldsymbol{I} \\ & +[1-\mu P_x]\rho E\left[\boldsymbol{r}(n)\boldsymbol{f}(\hat{h}^T(n))\right] \\ & +[1-\mu P_x]\rho E\left[\boldsymbol{f}(\hat{h}^T(n)\boldsymbol{r}(n))\right] \\ & +\rho^2 E\left[\boldsymbol{f}(\hat{h}(n))\boldsymbol{f}(\hat{h}^T(n))\right] + \mu^2 P_v P_x \boldsymbol{I}.\end{aligned} \tag{20}$$

Let $S(n) = tr[\boldsymbol{D}(n)]$, take the trace on both side of (20)

$$\begin{aligned}S(n+1) = {} & [1 - 2\mu P_x + (L+2)\mu^2 P_x^2] \cdot S(n) \\ & +2[1-\mu P_x]\rho c(n) \\ & +\rho^2 q(n) + L\mu^2 P_v P_x,\end{aligned} \tag{21}$$

where $c(n) = E\left[r^T(n)\boldsymbol{f}\left(\hat{h}(n)\right)\right], q(n) = \left\|\boldsymbol{f}(\hat{h}(n)\right\|_2^2$, $c(n)$ and $q(n)$ are all bounded and thus we can prove the condition of convergence as

$$|1 - 2\mu P_x + (\mathrm{L}+2)\mu^2 P_x| < 1.$$

Thereby we have:

$$0 < \mu < \frac{2}{(L+2)P_x}.$$

The final mean square deviation of reweighted l_{p}-norm penalty LMS is

$$S(\infty) = \frac{2[1-\mu P_x]\gamma c(\infty) + \gamma^2 \mu q(\infty) + L\mu P_v P_x}{P_x[2-(L+2)\mu P_x]}. \tag{21}$$

Dear Editors,

We would like to thank you and the reviewers for the valuable comments and suggestions on our manuscript (COM-2018-6186), which make it possible for us to improve our research works. Following your suggestion, we have revised the paper and accordingly updated the manuscript " Reweighted l_p Constraint LMS-Based Adaptive Sparse Channel Estimation for Cooperative Communication System ".

In the revision, we have carefully studied and addressed all the comments raised by the reviewers and revised the manuscript which we hope to meet the journal's standards for publication. The statement of the main corrections and the responses to the reviewer's comments are attached below. To ensure legibility, we typeset the comments in *italic*, our responses in **plain**, and rephrased sentences in the revised manuscript in **blue**. Note that all the citations, paragraphs, and pages referred herewith correspond to those appearing in the revised manuscript.

We hope that the manuscript has been improved in a satisfaction to you and the reviewers, and deeply appreciate your consideration of the revision.

Look forward to hearing from you.

Yours Sincerely,

Aihua Zhang

Pengcheng Liu

Bing Ning

Qiyu Zhou

Authors' Response to Reviewer 1

#1 Submitted by: Reviewer 1

1.*When comparing this version with the previous one, almost all figures have new performance curves. For example, for the Fig. 2, all the simulation setups are the same but the algorithm has approximately 10^-3 MSE for 100 iterations at 10 dB SNR and 10^-4 MSE for 100 iterations for 20 dB SNR for the previous version of the paper, however for the current version, they are approx. 7x10^-4 MSE for 230 iteration at 10 dB SNR and 2x10^-4 MSE for 200 iterations at 20 dB SNR. Is there anything different in the algorithm between two versions.*

We greatly thank the reviewer for this comment. We have selected new parameters for a better performance in the present manuscript, i.e., $\rho = 5 \times 10^{-4}$ and $\mu = 0.02$ in version 2 instead of $\rho = 3 \times 10^{-4}$ and $\mu = 0.05$ in version 1.

2. *It can be seen from Fig.3 and 5, when the sparsity, d, has high values (in here it is 4), the proposed algorithm has no significant error performance. How could you explain this?*

We greatly thank the reviewer for this comment. In this paper, the sparsity d denotes the number of non-zero coefficients. When d has high values, the channel will have more non-zero coefficients and the system will be less sparse. In this regard, the proposed algorithm has no significant performance.

3. *Still have typos and grammatical issues. i.e., h_i will be k_i in Section 2, page 6.*

We greatly thank the reviewer for the valuable comment to make our research more solid and convincing. We have corrected the typos, which are highlighted in blue color in the revised manuscript.

Authors' Response to Reviewer 2

The questions raised by the reviewer have been answered.

We greatly thank the reviewer for the helpful suggestions.