IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 8, 2019, accepted January 24, 2019, date of publication February 14, 2019, date of current version February 27, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897613

Distortion-Constraint-Based Group Sparse Channel Estimation Under α-Stable Noise

CHENGZHUO SHI, ZHENG DOU, (Member, IEEE), AND LIN QI
The College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

Corresponding author: Lin Qi (qilin@hrbeu.edu.cn)

This work was supported in part by the National Nature Science Foundation of China under Grant 61401115, Grant 61301096, and Grant 61671167, and in part by the International Exchange Program of Harbin Engineering University for Innovation-oriented Talents Cultivation.

ABSTRACT The ever-increasing requirements of wireless communications have inspired the search for a better method to tackle the problem of group sparse channel estimation in practical applications. Sparsity with group structure is encountered in numerous applications, but efforts to devise group sparse adaptive methods remain scarce, especially under impulse noise with symmetric alpha stable ($S\alpha S$) statistics. In this paper, we propose an improved adaptive algorithm using the distortion constraints based group sparse recursive least square (DC-GRLS) to exploit channel group sparsity and obtain robust performance under the background of α stable noise. We introduce distortion constraints combined with the mixed norms ($l_{p,q}$ norm), to obtain the relative balance between correctiveness and conservativeness. The MATLAB simulation results reveal that the improved algorithm can improve robustness under α stable noise when compared with the $l_{p,q}$ group algorithms and it can effectively predict the channel impulse response for a group sparse structure.

INDEX TERMS Group sparse structure, distortion constraints, mixed norms, symmetric alpha stable statistics, GRLS channel estimation.

I. INTRODUCTION

Group sparsity channels are typically encountered in a range of applications, such as source localization, cognitive spectrum sensing, and underwater acoustic and channel estimation applications [1], [2]. Energy distribution of the channel impulse response consists of blocks rather than a uniform distribution over the whole time delay domain [3], [4]; thus scholars have conducted extensive research on channel estimation algorithms for sparse channels in recent years [5], particularly to exploit the sparsity in channel prediction via kinds of adaptive filters [6]–[8]. Acquisition of steady-state offset and realize the channel tracking are essential to improving the estimation performance [9], [10], which is also applicable in wireless communication environments [11]. To improve sparse channel estimation performance in practical wireless communication systems as much as possible, development of the sparse recursive least squares (RLS) type channel estimation algorithms offers a potential solution [12]–[14]. This method has been proven as an effective approach. Zhu *et al.* [15] proposed an improved fast transversal filters for recursive least squares (FTRLS) filtering algorithm, which determined the amount of large error and accumulated it, then constructed error feedback to render the algorithm more stable. However, this algorithm is not suitable for sparse channel estimation. Chen *et al.* [16] proposed an RLS based fast adaptive sparse channel estimation algorithm, wherein two sparse constraint functions, L_1-norm and L_0-norm, are combined with different regularization parameters to obtain robust channel estimation. Eksioglu [3] presented a new analytic approximation for $l_{p,0}$ norm to utilize it as a group sparse regularizer, this was proven to be a better adaptive algorithm when a group sparse structure applies, but only Gaussian noise was considered in this paper. Zhu *et al.* [17] considered the RLS type based sparse channel estimation algorithms under a non-Gaussian noise background, where the non-Gaussian noise background followed a generalized Gaussian distribution (GGD) model. Pelekanakis and Chitre [18] extend the application of the algorithm to an α-stable noise background, and denoted two novel frameworks regarding how to design online adaptive algorithms. The first framework combined robust nonlinear methods with sparse-promoting L_0-norm regularization to obtain better sparse estimation performance; the second combined non-linear methods with a natural gradient (NG) to obtain better channel prediction. In this paper, sparse channel

The associate editor coordinating the review of this manuscript and approving it for publication was Jun Huang.

and non-Gaussian noise background are each considered. When applied to group sparse structure channel estimation, however the performance was not ideal.

The above papers, among many others, presented scholarly research on sparse channel estimation and group sparse channel estimation to obtain better adaptive estimation performance. Even so, efforts involving group sparse adaptive algorithms remain relatively scarce, particularly under non-Gaussian noise conditions. Recent work from Zimmermann; *et al.* showed that noise signals in practice (e.g., impulse noise), exhibit non-Gaussian behavior that obeys an α stable distribution [19]. Accordingly, more efficient channel estimation techniques will be needed in the future to ensure channel estimation under the practical noise.

In this paper, we propose a novel group sparse adaptive prediction algorithm by combining a distortion constraints with the mixed $l_{p,q}$ pseudo-norm as a group sparse regularizer, hereafter referred to as DC-distortion constraints based group sparse recursive least squares (GRLS). Inspired Su *et al.* [20], we believe that the quadratic complexity framework can maintain a balance between correctiveness and conservativeness [21], [22], effectively suppressing instability caused by impulse noise (i.e., frequent dramatic changes from $\hat{h}[n-1]$ to $\hat{h}[n]$ should decrease). We set the exponential approximation to the mixed $l_{p,0}$ pseudo-norm as part of the novel algorithm to promote group sparsity. In an attempt to develop the algorithm, we combine sub-gradients, including the $l_{\infty,1}$ and $l_{p,1}$ norm, with the distortion constraints for comparison. Numerous simulations suggest that our proposed sparse online estimation algorithm is effective for the group sparse channel under non-Gaussian noise.

The reminder of this paper is organized as follows. Section II introduces the system model, including the noise background model. Section III presents the novel algorithm and describes the derived the corresponding implementation process. The proposed algorithm is evaluated through simulations in Section IV, and we analyze the performance of the DC-GRLS algorithm on the basis of the experiments. Finally, Section V concludes the paper.

II. SYSTEM MODEL

In this paper, we consider a baseband channel representation of the channel with the linear input and output, as well as the additive noise process. First, we identify an actual channel impulse response for the communication system, expressed as $\mathbf{h}[n] = [h_0[n], h_1[n], \cdots, h_{N-1}[n]]^T$ at the series sample time n. The constructed $\mathbf{h}[n]$ is known with a group or block structure to the coefficient for $\hat{h}[n]$ as described above. Then the noisy output signal can be shown as:

$$y[n] = \mathbf{h}[n]^T\mathbf{x}[n] + w[n]. \tag{1}$$

where $\mathbf{x}[n] = [x[n], x[n-1], \cdots, x[n-N+1]]$ denotes the samples of the input signal at time n, and $w[n]$ denotes the additive noise affecting the system output. The purpose of adaptive channel estimation is to estimate the unknown channel vector $\mathbf{h}[n]$ using the known training signal $x[n]$ and received signal $y[n]$.

We assume that the actual baseband noise in the communication system is independent and identically distributed (i.i.d). Given the issues considered in this paper, we assume that $w[n]$ follows the symmetric alpha stable (SαS) distribution with a characteristic function as follows:

$$\varphi(w) = e^{-\gamma|w|^{\alpha}}. \tag{2}$$

where $\alpha \in (0, 2]$ is the characteristic exponent to describe the non-Gaussian degree, a smaller α corresponds to higher frequency pulses. Additionally, $\gamma > 0$ is the location parameter to describe the spread of the distribution. It should be note that when $\alpha = 2$, the probability density function (PDF) of the SαS distribution is equal to the Gaussian distribution, which reflects its versatility in describing actual system noise. For mathematical and practical reasons, we restrict the PDF to $\alpha \in (1, 2]$ for the SαS distribution; it is rare to find a condition for $\alpha < 1$ in practical communication systems [23], [24].

Generally, we analyze the performance of the algorithm in a communication system with different signal-to-noise ratio (SNR). When $\alpha \neq 2$ and $\alpha \neq 1$, the SαS noise does not have a closed form PDF; therefore, we must define the generalized signal to noise ratio (GSNR) for the channel estimation system:

$$GSNR(dB) = 10log_{10}\frac{\sigma_s p_s}{2\gamma^{\alpha}}. \tag{3}$$

where p_s is the signal power and σ_s is the ratio of the sample interval. $\gamma^{2/\alpha}$ denotes the noise dispersion, and GSNR $= E_s/N_0$ when $\alpha = 2$, which is consistent with Gaussian noise.

III. DISTORTION -CONSTRAINT -BASED MIXED-NORM GROUP SPARSE RLS ALGORITHM

In this section, we propose the DC-GRLS algorithm combined with distortion constraints; based on a general framework for online prediction algorithms proposed by Konstantinos Pelekanakis. We hope the adaptive estimation algorithm will be conservative and corrective when applied to a group sparse channel under non-Gaussian noise, thus the efficient cost function can be expressed as:

$$J[n] = f(\bar{e}[n]) + \delta D(\hat{h}[n], \hat{h}[n-1]). \tag{4}$$

where δ denote a non-negative regularization parameter to maintain a relative balance between correctiveness and conservativeness. Here, $f(\bar{e}[n])$ is the loss function, and $D(\cdot)$ is the distance function (i.e., the scalar distance function between $\hat{h}[n]$ and $\hat{h}[n-1]$). More importantly, this framework allows itself to incorporating additional constraints to obtained better prediction for $\hat{h}[n]$, and serving as the basis for many adaptive algorithms, as seen in [25]–[27]. Our main work is based on this framework.

A. IMPROVED CONSTRAINT FRAMEWORK

Based on the above mentioned algorithmic framework, distortion constraints for preventing drastic changes and group sparseness constraints can be integrated into a single objective function:

$$J[n] = \sum_{i=n-L+1}^{n} f(\bar{e}[i]) + \delta D(\hat{h}[n], \hat{h}[n-1]) + \zeta f(h). \quad (5)$$

Here

$$D(\hat{h}[n], \hat{h}[n-1]) = r[n]^T \mathbf{P}[n-1] r[n]. \quad (6)$$

$$f(h) = \| \mathbf{h} \|_{p,q} . \quad (7)$$

The first term is the loss function $f(\bar{e}[i])$ to ensure robustness against the outliers. ζ is the regularization parameter, and L (usually $L \leq 10$) is the length of the detection window (variable and limited by the channel and hardware requirements). To simplify the algorithm and yield less misadjustment in sparse channels, we select $f(\bar{e}[i]) =| \bar{e}[i] |_1$ as the new loss function to accommodate the large impulse.

The matrix $\mathbf{P}[n]$ is a $K \times K$ Riemannian metric tensor, it is a self-adjoint and positive matrix. The term is

$$D(\hat{h}[n], \hat{h}[n-1]) = r[n]^T \mathbf{P}[n-1] r[n]. \quad (8)$$

We use it to represent the Riemannian distance between $\hat{h}[n]$ and $\hat{h}[n-1]$. According to the description in [36], the ordinary Euclidean gradient does not have the fastest convergence speed in Riemannian spaces, hence why regularizing $J[n]$ with a Riemannian is a better choice to accelerate convergence of the target coefficient.

We must also determine how to select the proper $\mathbf{P}[n]$ for the channel with group sparse structure. We refer to the scheme in [28] and [29], where the parameter space of a sparse channel may be visualized as a space with wrapping: the Euclidean distance is smaller than any distance in the direction orthogonal to the coordinate axis. Then we set $\mathbf{P}[n]^{-1} = \mathbf{G}[n]$, where $\mathbf{G}[n]$ is a proportionate diagonal matrix with elements $g_k[n]_{k=0}^{K-1}$, and calculated as follows:

$$g_k[n] = \frac{1-\beta}{2K} + (1+\beta)\frac{| \hat{h}_k[n] |_1}{2 \| \hat{\mathbf{h}} \|_1 + \varepsilon}. \quad (9)$$

where ε should be a positive and small constant to make sure that the algorithm does not run correctly during the iteration. β should be chosen according to the sparseness of the channel, where $\beta \in (0, 0.5)$ indicates that channel is very sparse. δ can be obtained through $\delta = (1-\beta)\delta'/2K$ in ref [7].

$\zeta f(h)$ is an incorporated additional constraint to develop group structured algorithms for block-sparsity aware adaptive filtering. We assume the group structure is priori known in this paper. As shown in (10), $h_{g_i} \in \mathbf{R}^N$, the new expression for the group structure of vectors is as follows:

$$\{h_{g_i}\}_k = \begin{cases} h_k, & if\ k \in g_i; \\ 0, & if\ k \notin g_i. \end{cases} \quad (10)$$

where $k = 0, \cdots, N-1$. We set the channel by zeroing all values of the coefficient vector $\mathbf{h}$ except the positions where h_{g_i} located. To develop new group sparsity cognizant GRLS algorithms, we define G as the total number of groups and with no overlap between each other:

$$g_i \cap g_j = \oslash, \quad i \neq j,\ g_i,\ g_j \in G. \quad (11)$$

Next, we give the sub-gradient of mixed norms and induce them as regularizing penalty functions into cost function (5).

B. GROUP SPARSITY AND MIXED NORMS

The mixed $l_{p,q}$ norm corresponding to the proposed group channel is defined as:

$$\begin{aligned} \| \mathbf{h} \|_{p,q} &= (\sum_{i=1}^{G}(\| h_{g_i} \|_p)^q)^{-q} \\ &= (\sum_{i=1}^{G}(\sum_{k \in g_i k_k^p}^{q} /p))^{1/q}. \end{aligned} \quad (12)$$

Eq. (12) is convex when $p, q \in [1, \infty]$; However, the expression for actual group sparsity is the mixed pseudo-norm, which can be described as

$$\| \mathbf{h} \|_{p,0} = \sum_{i=1}^{G} I(\| h_{g_i} \|_p). \quad (13)$$

The structure presented in (13) is non-convex and not applicable to the sub-gradient analysis. Therefore, we adopt two special mixed norms as promoters of group sparsity. One is $l_{2,1}$ that utilized in [16]

$$\| \mathbf{h} \|_{2,1} = \sum_{i=1}^{G} \| \mathbf{h}_{g_i} \|_2 = \sum_{i=1}^{G} (\sum_{k \in g_i} h_k^2)^{1/2}. \quad (14)$$

The other is a novel approximation mixed norm based on $l_{p,0}$

$$\| \mathbf{h} \|_{p,0} \approx \sum_{i=1}^{G} \| \mathbf{h} \|_{p,0^\eta} = G - \sum_{i=1}^{G} e^{-\eta \| \mathbf{h}_{g_i} \|_p}. \quad (15)$$

where η is a small and positive constant. The sub-gradients of these functions will be calculated as new constraint in the following sections to obtain better performance in group sparse channel estimation.

C. DISTORTION CONSTRAINT BASED GROUP SPARSE RLS ALGORITHM

The upgraded equation of the channel estimation algorithm can be derived from $\nabla_{r[n]} J[n] = 0$, which can be calculated as

$$\begin{aligned} \nabla_{r[n]} J[n] = \nabla_{r[n]} (&\sum_{i=n-L+1}^{n} f(\bar{e}[i])) \\ &+ \nabla_{r[n]} (r[n]^T \mathbf{P}[n-1] r[n]) \\ &+ \nabla_{r[n]} (\zeta f(h)). \end{aligned} \quad (16)$$

where $J[n]$ is given by (5), and each of the above terms can be computed individually as

$$\begin{aligned}
&\nabla_{r[n]}(\sum_{i=n-L+1}^{n} f(\bar{e}[i])) \\
&= -\sum_{i=n-L+1}^{n} \psi(\bar{e}[i])\mathbf{x}[i] \\
&= -\sum_{i=n-L+1}^{n} q(\bar{e}[i])\bar{e}[i]\mathbf{x}[i] \\
&= -\sum_{i=n-L+1}^{n} q(\bar{e}[i])(\bar{e}[i] - \mathbf{x}[i]^T r[n])\mathbf{x}[i] \\
&= -\sum_{i=n-L+1}^{n} q(\bar{e}[i])\bar{e}[i]\mathbf{x}[i] \\
&\quad + \sum_{i=n-L+1}^{n} q(\bar{e}[i])(\mathbf{x}[i]\mathbf{x}[i]^T)r[n] \\
&= -X[n]Q[n]e[n] + X[n]Q[n]X[n]^T r[n]. \qquad (17)
\end{aligned}$$

where $Q[n]$ is an L*L diagonal matrix with elements $q(e[i])$. For the second term we have:

$$\nabla_{r[n]}(\delta r[n]^T \mathbf{P}[n-1]r[n]) = \delta \mathbf{P}[n-1]r[n]. \qquad (18)$$

$\nabla_{r[n]}(\zeta f(h))$ is computed as follows:

1)

if $f(\mathbf{h}) = \| \mathbf{h} \|_{2,1}$, the sub-gradients $\nabla_{r[n]}(\zeta f(h))$ can be calculated as

$$\begin{aligned}
\nabla_{r[n]}(\zeta f(h)) &= \sum_{i=1}^{G} \zeta \nabla_{r[n]} \| \mathbf{h}_{g_i} \|_2 \\
&\approx \sum_{i=1}^{G} \frac{\zeta \mathbf{h}_{g_i}}{\| \mathbf{h}_{g_i} \|_2 + \sigma}. \qquad (19)
\end{aligned}$$

where we use sub-gradient approximation to simplify the calculation progress, and σ is again a small and positive constant. To calculate of the iteration, we present V[n] as

$$V[n] = \sum_{i=1}^{G} \frac{\mathbf{h}_{g_i}}{\| \mathbf{h}_{g_i} \|_2 + \sigma}. \qquad (20)$$

2) if $f(\mathbf{h}) = \| \mathbf{h} \|_{p,0}$, the sub-gradients $\nabla_{r[n]}(\zeta f(h))$ can be calculated as

$$\nabla_{r[n]}(\zeta f(h)) = \sum_{i=1}^{G} \zeta \eta \nabla_{r[n]}\{\| \mathbf{h}_{g_i} \|_p\} e^{-\eta \|\mathbf{h}_{g_i}\|_p}. \qquad (21)$$

Similarly, to simplify this algorithm, it can be expressed as $e^{-\eta|x|} \approx (1 - \eta \mid x \mid)_+$ when $\mid x \mid \leq 1/\eta$. Thus, (21) can be expressed more succinctly as

$$\nabla_{r[n]}(\zeta f(h)) = \sum_{i=1}^{G} \zeta \eta \nabla_{r[n]}\{\| \mathbf{h}_{g_i} \|_p\}(1 - \eta \| \mathbf{h}_{g_i} \|_p)_+. \qquad (22)$$

$(f(x))_+$ indicates the maximum value of (0, f(x)). According to research on the algorithm, only the cases of $p = 1$ and $p = 2$ are considered. When $p = 1$, the matching sub-gradient can be presented as

$$\nabla_{r[n]}(\zeta f(h)) = \sum_{i=1}^{G} \zeta \eta sgn(\mathbf{h}_{g_i})(1 - \eta \| \mathbf{h}_{g_i} \|_1)_+. \qquad (23)$$

where the corresponding V[n] is

$$V[n] = \sum_{i=1}^{G} \eta sgn(\mathbf{h}_{g_i})(1 - \eta \| \mathbf{h}_{g_i} \|_1)_+. \qquad (24)$$

When $p = 2$, the corresponding sub-gradient can be computed as

$$\begin{aligned}
\nabla_{r[n]}(\zeta f(h)) &= \sum_{i=1}^{G} \zeta \eta \nabla_{r[n]}\{\| \mathbf{h}_{g_i} \|_p\}(1 - \eta \| \mathbf{h}_{g_i} \|_2)_+ \\
&\approx \sum_{i=1}^{G} \frac{\zeta \eta}{\| \mathbf{h}_{g_i} \|_2 + \sigma}(1 - \eta \| \mathbf{h}_{g_i} \|_2)_+. \qquad (25)
\end{aligned}$$

where the corresponding V[n] is

$$V[n] = \sum_{i=1}^{G} \frac{\eta}{\| \mathbf{h}_{g_i} \|_2 + \sigma}(1 - \eta \| \mathbf{h}_{g_i} \|_2)_+. \qquad (26)$$

D. CHANNEL UPDATE EQUATION FOR DC-GRLS

The derivation process of the novel algorithm was specifically described in the previous section. The following equations represent the main process for group sparse channel estimation with distortion constraints.

By setting $\nabla_{r[n]} J[n] = 0$, and combining(17),(18), and (19) (or (22)), we obtain

$$\begin{aligned}
&X[n]Q[n]e[n]^* \\
&\quad = r[n](\delta P[n-1] + X[n]Q[n]X[n]^*) + \zeta v[n]. \qquad (27) \\
r[n] &= (\delta P[n-1] + X[n]Q[n]X[n]^T)^{-1} \\
&\qquad (x[n]Q[n]e[n] - \zeta V[n]) \\
&= (\delta P[n-1] + X[n]Q[n]X[n]^T)^{-1} x[n]Q[n]e[n] \\
&\quad - (\delta P[n-1] + X[n]Q[n]X[n]^T)^{-1} \zeta V[n]. \qquad (28)
\end{aligned}$$

In this equation, we aim to solve the problem related to non-linear impulse $h[n]$, which determins the $V[n]$ and $r[n]$. To simplify the calculation process for $r[n]$, we assume that $V[n] \simeq V[n-1]$; then, combined with the matrix inversion theorem, we obtain

$$\begin{aligned}
r[n] = &\frac{k[n]}{X[n]^T k[n] + \delta Q[n]^{-1}} e[n] \\
&- \frac{\zeta}{\delta}(P[n-1]^{-1} - K[n]X[n]^T \\
&\times (P[n-1]^{-1})^T)V[n-1]. \qquad (29)
\end{aligned}$$

where $r[n] = \hat{h}[n] - \hat{h}[n-1]$. Next, we have the update process for prediction algorithms (taking the $l_{1,0}$ norm as an example)

$$\begin{aligned}
e[n] &= y[n] - X[n]^T \hat{h}[n-1], \\
k[n] &= P[n-1]^{-1} X[n],
\end{aligned}$$

$$K[n] = \frac{k[n]}{X[n]^T k[n] + \delta Q[n]^{-1}},$$
$$P[n]^{-1} = P[n-1]^{-1} - K[n]X[n]^T (P[n-1]^{-1})^T,$$
$$V_k[n] = \sum_{k=1}^{K} \eta sgn(\mathbf{h}_{g_k})(1 - \eta \parallel \mathbf{h}_{g_k} \parallel_1)_+,$$
$$\hat{h}[n] = \hat{h}[n-1] + \mu K[n]e[n] - \frac{\zeta\mu}{\delta} P[n]^{-1} V[n-1].$$

where $\mu \in (0, 1]$ is the step-length parameter. $X[n]$, $P[n]$ and $Q[n]$ have been defined in previous section. Initialization of the algorithm starts with $\hat{h}[0] = 0$.

IV. SIMULATION RESULTS

We have presented theoretical analysis of the novel algorithm for group sparse channel estimation under non-Gaussian noise. In this section, we test the performance of the proposed algorithm in a group of sparse channels with known prior information through numerous simulations.

A. NOTATIONS

Observation noise follows the SαS distribution, and the characteristic $\alpha \in (1.5, 2]$ will be more consistent with the measured results [30]. The SNR is defined as Eq (3). The system impulse response $\mathbf{h}$ has total 64 coefficients including $G = 16$ blocks with 4 coefficients per block. N is the number of non-zero blocks that can be selected for performance tests of the proposed algorithms. The typical group sparse channel we constructed with 4 non-zero blocks is shown in Fig. 1.

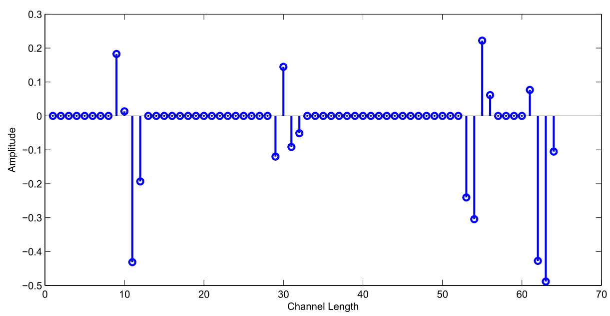


FIGURE 1. Group sparse channel with 4 non-zero groups of 64 coefficients.

The simulation parameters are as follows:

$\beta = 0.5$ indicates the sparseness of the channel, $L = 4$ is the length of the observation window, $\delta' = 10$ when calculating $\delta, \zeta = 5 \cdot 10^{-4}$ denotes the regulation parameters, $\eta = 10$ is a constant and $\mu = 0.1$ for the updated step-size. The performance measure is the average mean squared deviation (MSD), defined as $MSD = 10log(\parallel \mathbf{h} - \hat{\mathbf{h}} \parallel_2^2 / \parallel \mathbf{h} \parallel_2^2)$.

As shown in Fig. 2, the received signal with additive non-Gaussian noise was applied in the simulations to test the algorithm's performance.

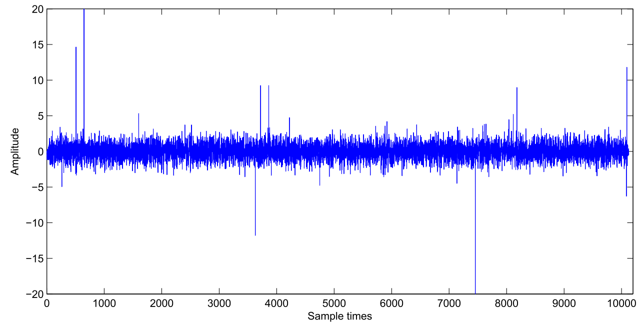


FIGURE 2. Received signal with additive α stable noise, $SNR = 10$ dB, $\alpha = 1.6$.

B. SIMULATION RESULTS

The channel to be estimated with α-stable noise was described in the previous section. Here, we will compare the performance of the algorithms with different SNR and different noise characteristic exponents. We select $l_{2,0}$-GRLS and basic RLS algorithms for comparison, whereas RLSA is the basic algorithm and $l_{2,0}$-GRLS algorithm perform best when applied to group sparse channel estimation under Gaussian noise.

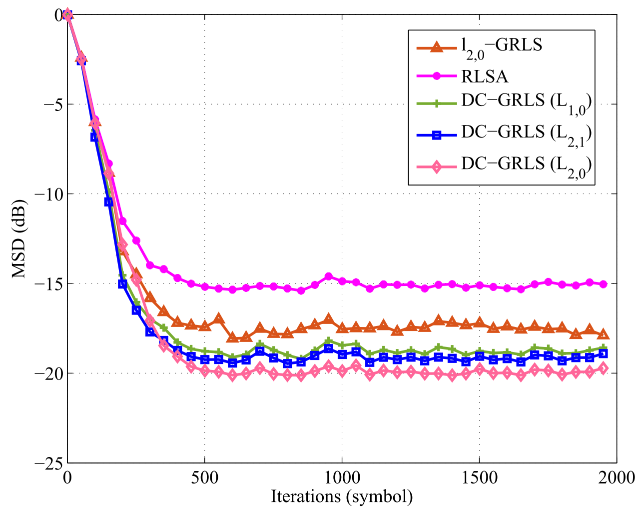


FIGURE 3. Prediction curves of DC-GRLS algorithms for different α and SNR, and the comparison with the traditional GRLS algorithms ($\alpha = 1.5$, $SNR = 10$ dB, N = 2).

As indicated in Fig. 3, characteristic exponents $\alpha = 1.5$, signal to noise ratio $SNR = 10$ dB, and the non-zeros groups N = 2. We considered the DC-GRLS algorithms with $l_{1,0}$, $l_{2,0}$, and $l_{2,1}$ norms and compared them with the traditional GRLS prediction algorithms. The prediction of DC-GRLS($l_{2,0}$) algorithm returned lower MSD than traditional GRLS algorithms while the DC-GRLS algorithms with ($l_{1,0}$) and (($l_{2,1}$)) are very similar to each other.

Then, we set characteristic exponents $\alpha = 1.7$, and SNR = 10 dB. Prediction results are depicted in Fig. 4. The DC-GRLS($l_{2,0}$) algorithm was consistently robust and exhibited the lowest MSD for $\alpha = 1.7$ whereas the DC-GRLS($l_{1,0}$)

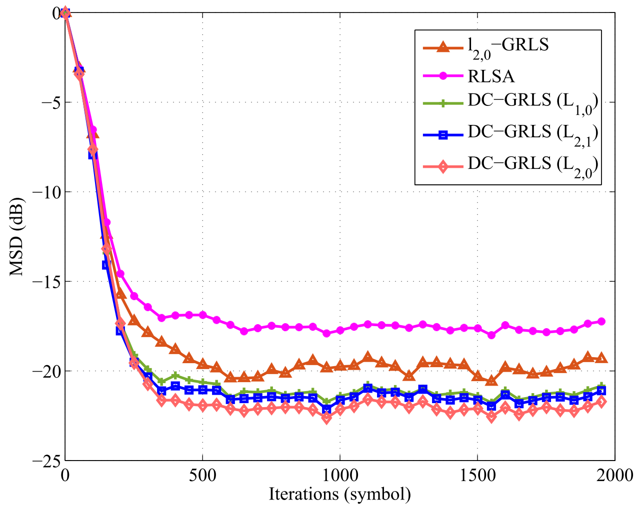


FIGURE 4. Prediction curves of DC-GRLS algorithms for different α and SNR, and the comparison with the traditional GRLS algorithms ($\alpha = 1.7$, SNR = 10 dB, non-zero blocks N = 2).

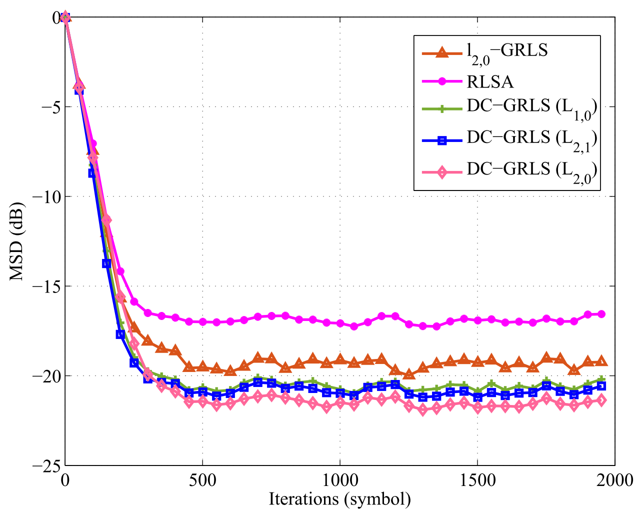


FIGURE 5. Prediction curves of DC-GRLS algorithms and the comparison with the traditional GRLS algorithms ($\alpha = 1.5$, SNR = 20 dB, non-zero blocks N = 2).

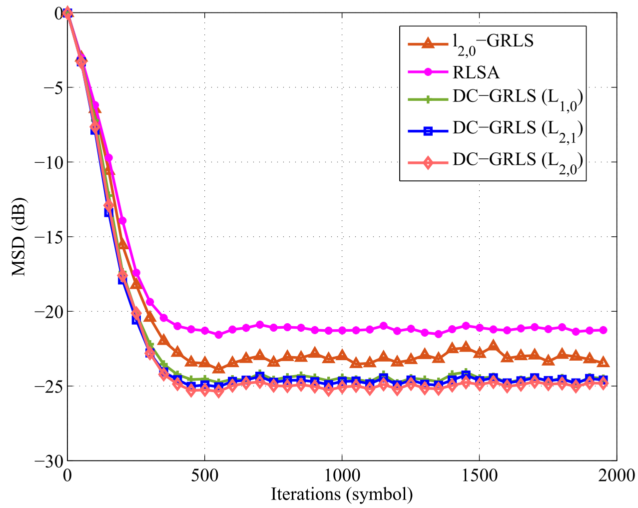


FIGURE 6. Prediction curves of DC-GRLS algorithms and the comparison with the traditional GRLS algorithms ($\alpha = 1.7$, SNR = 20 dB, non-zero blocks N = 2).

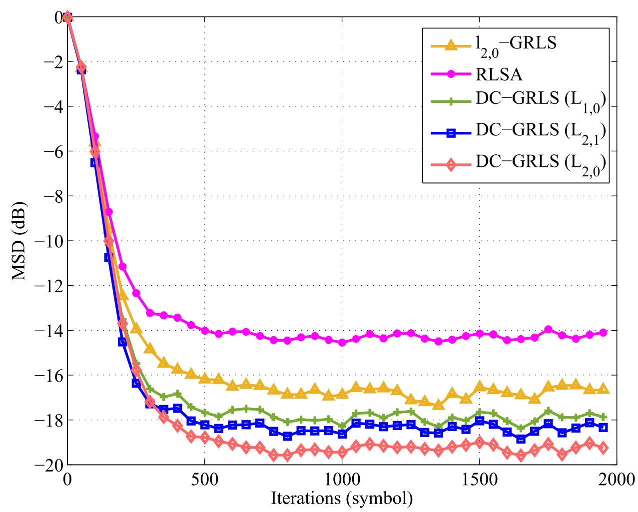


FIGURE 7. Prediction curves of DC-GRLS algorithms and the comparison with the traditional GRLS algorithms; for actual noise with SNR $\approx$ 10 dB, non-zero blocks of priori known channel N = 2.

and DC-GRLS($l_{2,1}$) algorithms did not show superiority obviously.

Comparisons with higher SNR are shown in Fig. 5 and Fig. 6. We set SNR $=$ 20 dB and α $=$ 1.5 and 1.7 for simulations to compare the algorithm performance. The DC-GRLS($l_{2,0}$) algorithm was consistently robust and have a faster convergence rate when SNR $=$ 20 dB; therefore, simulations have verified the effectiveness of the proposed DC-GRLS algorithm.

To further test the performance of the proposed algorithm, we conducted simulations by superimposing actual noise on the group sparse channel for comparison.

In the simulations, we could not obtain an accurate characteristic exponent of the actual noise; we could only set the SNR $\approx$ 10 dB. Simulation result are shown in Fig. 7.

Algorithm performance declined under actual noise. The RLS algorithm could not obtain a good performance when cope with the prediction under actual noise. We also find that The prediction process of $l_{2,0}$-GRLS algorithm was unstable, while DC-GRLS($l_{2,0}$) still shows the superiority of other algorithms. Thus it proves that the novel algorithm we proposed in this paper is practical. For comparison, we increased the signal to noise ratio to SNR $\approx$ 20 dB, and the prediction MSD curves are shown in Fig. 8.

The proposed algorithm also showed good adaptability when we increased the SNR. The convergence speed was also faster than with a lower SNR, and the MSD was close to the simulation results. The performance of channel estimation algorithms improved when increased the SNR, and DC-GRLS($l_{2,0}$) still shows the superiority of other algorithms.

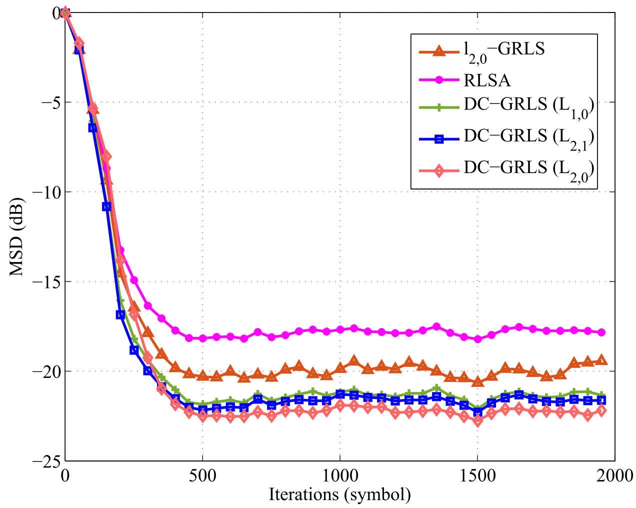


FIGURE 8. Prediction curves of DC-GRLS algorithms and the comparison with the traditional GRLS algorithms, SNR ≈ 20 dB, non-zero blocks of priori known channel N = 2.

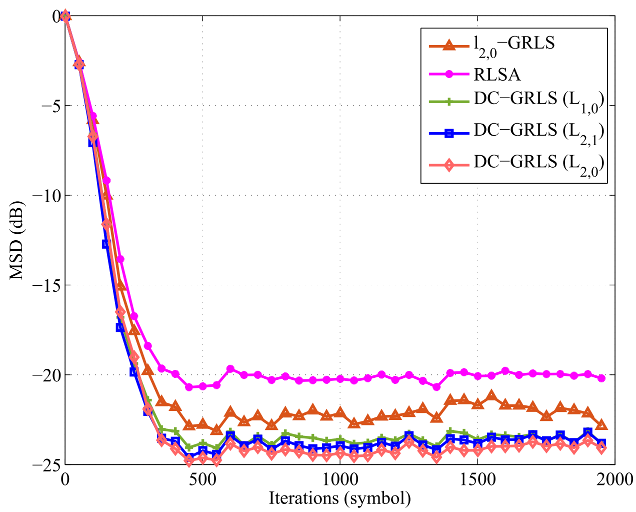


FIGURE 9. Prediction curves of DC-GRLS algorithms and the comparison with the traditional GRLS algorithms, SNR ≈ 20 dB, non-zero blocks of priori known channel N = 3.

We then set the non-zero blocks of the channel N = 3 (the channel contained priori known 128 coefficients). The noise captured the actual noise sequence, and SNR ≈ 10 dB; other simulation parameters for the proposed algorithm were described in the notation section. Fig. 9 illustrates the MSD curves when N = 3 for the group sparse channel.

Channel information and the prediction results are listed in Table.1 for comparison. The real values are the coefficients of the priori channel we constructed, and estimation results are the prediction results. The predicted results approximated the actual value, indicating that the proposed algorithm can accurately predict the group sparse channel impulse response.

TABLE 1. Comparison of the coefficients of the priori channel and corresponding prediction results of the channel.

Coefficients	1	2	3	4
Real values	-0.4	0.16	-5.8	-0.41
Estimation	-0,41	0.58	-5.79	-0.4
Real values	0.03	0.16	-0.32	0.07
Estimation	0.02	0.16	-0.315	0.065
Real values	0.01	-0.19	-0.25	0.28
Estimation	0	-0.17	-0.45	0.28

V. CONCLUSION

A novel adaptive group sparse channel estimation algorithms has been introduced in this paper. A prediction algorithm based on a combination of distortion constraint and mixed norm was derived. The focus of this paper was on improving algorithm performance in the presence of SαS noise. The distortion constraint maintained a relative balance between correctiveness and conservativeness, induced by the loss function, whereas the mixed-norms were a good choice to promote the group sparsity. Fitted noise and the actual noise were each considered in the experiment. Simulations demonstrated that our adaptive algorithm is more robust and accurate when predicting the group sparse channel under SαS noise, thus outperforming previously developed systems.

REFERENCES

[1] G. Gui, W. Peng, and F. Adachi, "Improved adaptive sparse channel estimation based on the least mean square algorithm," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 3105–3109.
[2] H. Mashud and M. Kaushik, "Direction-of-arrival estimation using a mixed$\ell_{2,0}$norm approximation," *IEEE Trans. Signal Process*, vol. 58, no. 9, pp. 4646–4655, Sep. 2010.
[3] E. M. Eksioglu, "Group sparse RLS algorithms," *Int. J. Adapt. Control Signal Process.*, vol. 28, no. 12, pp. 1398–1412, Dec. 2014.
[4] G. Gui, L. Xu, and F. Adachi, "RZA-NLMF algorithm-based adaptive sparse sensing for realizing compressive sensing," *EURASIP J. Adv. Signal Process.*, vol. 1, p. 125, Dec. 2014.
[5] Y. Tu *et al.*, "Semi-supervised learning with generative adversarial networks on digital signal modulation classification," *CMC-Comput. Mater. Continua*, vol. 55, no. 2, pp. 243–254, 2018.
[6] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.
[7] Y. Gu, J. Jin, and S. Mei, "ℓ_0 norm constraint LMS algorithm for sparse system identification," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 774–777, Sep. 2009.
[8] G. Gui, W. Peng, and F. Adachi, "Adaptive system identification using robust LMS/F algorithm," *Int. J. Commun. Syst.*, vol. 27, no. 11, pp. 2956–2963, Nov. 2014.
[9] Z. Zhang, X. Guo, and Y. Lin, "Trust management method of D2D communication based on RF fingerprint identification," *IEEE Access*, vol. 6, pp. 66082–66087, 2018.
[10] D. Angelosante, J. A. Bazerque, and G. B. Giannakis, "Online adaptive estimation of sparse signals: Where RLS meets the ℓ_1-norm," *IEEE Trans. signal Process.*, vol. 58, no. 7, pp. 3436–3447, Jul. 2010.
[11] D. Méndez-Romero and M. J. F. G. Garcia, "Simpler multipath detection for vehicular OFDM channel tracking," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10752–10759, Nov. 2018.
[12] C. Yang *et al.*, "Optimization or alignment: Secure primary transmission assisted by secondary networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 905–917, Apr. 2018.

[13] M. Jia, Z. Yin, Q. Guo, G. Liu, and X. Gu, "Waveform design of zero head DFT spread spectral efficient frequency division multiplexing," *IEEE Access*, vol. 5, pp. 16944–16952, 2017.

[14] G. Ding, Q. Wu, Y.-D. Yao, J. Wang, and Y. Chen, "Kernel-based learning for statistical signal processing in cognitive radio networks: Theoretical foundations, example applications, and future directions," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 126–136, Jul. 2013.

[15] J. Zhu, J. Zhang, and Q. Chen, "An improved FTRLS filtering algorithm and its simulation analysis," in *Proc. IEEE 17th Int. Conf. Commun. Technol. (ICCT)*, Oct. 2017, pp. 1709–1714.

[16] Y. Chen, G. Gui, and L. Wang, "Who care for channel sparsity? Robust sparse recursive least square based channel estimation," in *Proc. 8th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2016, pp. 1–5.

[17] X. Zhu *et al.*, "Recursive least square based sparse channel estimation under non Gaussian noise background," *Appl. Electron. Techn.*, vol. 42, no. 6, pp. 109–112, 2016.

[18] K. Pelekanakis and M. Chitre, "Adaptive sparse channel estimation under symmetric alpha-stable noise," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3183–3195, Jun. 2014.

[19] M. Zimmermann and K. Dostert, "Analysis and modeling of impulsive noise in broad-band powerline communications," *IEEE Trans. Electromagn. Compat.*, vol. 44, no. 1, pp. 249–258, Feb. 2002.

[20] G. Su, J. Jin, Y. Gu, and J. Wang, "Performance analysis of ℓ_0 norm constraint least mean square algorithm," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2223–2235, May 2012.

[21] R. E. Mahony and R. C. Williamson, "Prior knowledge and preferential structures in gradient descent learning algorithms," *J. Mach. Learn. Res.*, vol. 1, pp. 311–355, Sep. 2001.

[22] S. L. Gay and S. C. Douglas, "Normalized natural gradient adaptive filtering for sparse and non-sparse systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2002, pp. II-1405–II-1408.

[23] A. T. Georgiadis and B. Mulgrew, "Adaptive Bayesian decision feedback equaliser for alpha-stable noise environments," *Signal Process.*, vol. 81, no. 8, pp. 1603–1623, Aug. 2001.

[24] G. Gui, L. Xu, W. Ma, and B. Chen, "Robust adaptive sparse channel estimation in the presence of impulsive noises," in *Proc. IEEE Int. Conf. Digit. Signal Process. (DSP)*, Jul. 2015, pp. 628–632.

[25] L. R. Vega, H. Rey, J. Benesty, and S. Tressens, "A family of robust algorithms exploiting sparsity in adaptive filters," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 4, pp. 572–581, May 2009.

[26] T. Shao, Y. R. Zheng, and J. Benesty, "An affine projection sign algorithm robust against impulsive interferences," *IEEE Signal Process. Lett.*, vol. 17. no. 4, pp. 327–330, Apr. 2010.

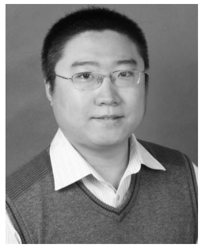
[27] K. Pelekanakis and M. Chitre, "A class of affine projection filters that exploit sparseness under symmetric alpha-stable noise," in *Proc. MTS/IEEE OCEANS*, Jun. 2013, pp. 1–5.

[28] D. L. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. Speech Audio Process.*, vol. 8, no. 5, pp. 508–518, Sep. 2000.

[29] J. Benesty and S. L. Gay, "An improved PNLMS algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Orlando, FL, USA, May 2002, pp. 1881–1884.

[30] C. Shi, S. Zheng, Y. Lin, and W. Li, "Dynamic threshold-setting for RF-powered cognitive radio networks in non-Gaussian noise," *Phys. Commun.*, vol. 27, pp. 99–105, Apr. 2018.

CHENGZHUO SHI received the B.S. degree in electrical information from Harbin Engineering University, Harbin, China, in 2013, where he is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering. His current research interests include cognitive radio networks, intelligent communication systems, and control signal processing.



ZHENG DOU received the B.S. degree from Harbin Engineering University, China, in 2001, and the Ph.D. degree from the College of Information and Communication Engineering, Harbin Engineering University, in 2007, where he is currently a Professor. Until now, he has presented more than 60 research papers, 20 patents, and three books. His research interests include intelligent wireless communication, cognitive radio, and the design of communication signal waveform. He has been a Senior Member of the China Communications Society. He has served as a Reviewer for several journals and conferences such as IEEE Communications, *Signal processing*, the IEEE ACCESS, the *Journal of Communications*, and *Ship Engineering*. He has been a Reviewer of the National Natural Science Foundation.

LIN QI received the B.S. degree in communication engineering, the M.S. degree in signal and information processing, and the Ph.D. degree in communication and information system from Harbin Engineering University, China, in 2002, 2005, and 2011, respectively, where she was an Assistant Professor with the College of Information and Communication Engineering, from 2002 to 2005, and has been a Lecturer, since 2007. From 2015 to 2016, she was with the Institute of Materials and Systems for Sustainability, Nagoya University, as a Visiting Scholar. Her current research interests include wideband digital communication systems, SDR, and communication signal processing.