

# An Overloaded MU-MIMO Signal Detection Method Using Piecewise Continuous Nonconvex Sparse Regularizer

Atsuya Hirayama\* and Kazunori Hayashi†

\* Graduate School of Engineering, Osaka City University, 3-3-138 Sugimoto, Sumiyoshi-ku, Osaka, 558-8585 Japan  
E-mail: m20tb035@ka.osaka-cu.ac.jp

† Center for Innovative Research and Education in Data Science / Graduate School of Informatics, Kyoto University, 301 Konoekan, 69, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8315 Japan  
E-mail: hayashi.kazunori.4w@kyoto-u.ac.jp

**Abstract**—In this paper, we consider the signal detection problem of overloaded massive multi-user multi-input multi-output (MU-MIMO) orthogonal frequency division multiplexing (OFDM) and single carrier block transmission with cyclic prefix (SC-CP) systems. For the systems, we employ iterative weighted sum of complex sparse regularizers with group sparsity (IWSCSR-GS) optimization, which is a complex discrete-valued vector reconstruction method that uses discreteness of symbols to estimate unknown vectors, and propose a signal reconstruction method using piecewise continuous nonconvex sparse regularizers, such as smoothly clipped absolute deviation (SCAD) or minimax concave penalty (MCP), in the optimization problem. Computer simulation results demonstrate that the proposed signal reconstruction method with MCP achieves better symbol error rate (SER) performance than that of not only IWSCSR-GS with  $\ell_1$  norm but also that with  $\ell_p$  norm ( $p = 0, 1/2, 2/3$ ) or  $\ell_1 - \ell_2$  difference, which are nonconvex sparse regularizers, and the proposed signal reconstruction method with SCAD achieves the best performance among the methods for large systems with high signal-to-noise ratio (SNR) region.

## I. INTRODUCTION

With the advent of the 5th generation mobile communications system (5G), the requirement for simultaneous connections of massive terminals has become an important aspect of wireless communications. This requirement is expected to accelerate as wireless systems evolve toward the next generation, i.e., 6G [1]–[3]. One of the technologies supporting such systems is massive multi-user multi-input multi-output (MU-MIMO), because the data collection from a large number of internet-of-things (IoT) terminals at a base station with multiple antennas can be modeled by the MU-MIMO system. In common MIMO systems, including massive MU-MIMO systems, the number of transmit antennas (streams) is assumed to be less than or equal to that of receive antennas. However, in typical IoT environments, the number of transmit terminals could be greater than that of receive antennas even when a massive antenna array is employed at the base station. Therefore, it is difficult to apply conventional MIMO signal detection methods in such environments.

MIMO communications where the number of transmit streams is greater than that of receive antennas are called

overloaded MIMO, and the signal detection problem in such cases is very difficult because it is underdetermined. Maximum likelihood (ML) approach can obtain the estimate of the transmitted signals if a finite alphabet is used for the transmitted symbols as in digital communications [4], but the ML approach is prohibitively computationally expensive and infeasible for massive overloaded MIMO communications. Therefore, we have proposed a low computational complexity MIMO signal detection scheme [5] using convex optimization, where sum-of-absolute-values (SOAV) optimization [6], which is based on the idea of compressed sensing [7], [8], is employed. Moreover, by extending the method from the real domain to the complex domain, we have proposed a sparse complex discrete-valued vector reconstruction method named sum of complex sparse regularizers (SCSR) optimization [9] in [10]. Furthermore, noting that when orthogonal frequency division multiplexing (OFDM) or single carrier block transmission with cyclic prefix (SC-CP) is used in the IoT environments, the signals to be estimated have not only the discreteness but also the group sparsity, we have proposed SCSR with group sparsity (SCSR-GS) [11], which can take group sparsity into consideration. In addition, we have extended the SCSR-GS optimization into an iterative approach named iterative weighted SCSR-GS (IWSCSR-GS) [11], where we iteratively solve the SCSR-GS optimization problem with updating the parameters in the objective function. Most recently, in [11], we have improved the signal detection performance of IWSCSR-GS by using nonconvex sparse regularization function, such as  $\ell_p$  norm ( $p = 0, 1/2, 2/3$ ) and  $\ell_1 - \ell_2$  difference, for the discrete regularization.

In this paper, in order to further improve the performance of IWSCSR-GS, we focus on smoothly clipped absolute deviation (SCAD) and minimax concave penalty (MCP), which are nonconvex sparse regularization functions that have been shown to have better performance than  $\ell_1$  norm regularization [12] for compressed sensing, and propose a method to use them for discrete regularizers in IWSCSR-GS. These regularizers are piecewise continuous nonconvex functions whose behavior depends on several variables, which we call the

nonconvexity parameters. They include  $\ell_1$  norm regularization as a special case when they are set as infinity. The proximal operators of the regularizers can be computed in real space, and we use them in the alternating direction method of multipliers (ADMM) [13], [14] based algorithm to solve IWSCSR-GS optimization problem. Computer simulation results show that the proposed IWSCSR-GS with MCP gives better symbol error rate (SER) performance than our conventional method with  $\ell_p$  norm ( $p = 0, 1, 1/2, 2/3$ ) or  $\ell_1 - \ell_2$  difference and that with SCAD achieves the best performance among the methods for large systems with high signal-to-noise ratio (SNR) region.

In the rest of the paper, the following notations are used.  $\mathbb{R}$  is the set of all real numbers and  $\mathbb{C}$  is the set of all complex numbers. Let  $\mathbf{0}_Q$  and  $\mathbf{1}_Q$  denote a vector of size  $Q$  with all 0 elements and with all 1 elements, respectively.  $(\cdot)^T$  and  $(\cdot)^H$  are the transpose and Hermitian transpose, respectively.  $\mathbf{I}_N$  is an identity matrix of size  $N \times N$ . For a complex vector  $\mathbf{u} = [u_1 \cdots u_N]^T \in \mathbb{C}^N$  and an operator  $f(\cdot) : \mathbb{C}^N \rightarrow \mathbb{C}^N$ ,  $[f(\mathbf{u})]_n$  and  $u_n$  represent the  $n$ -th components of  $f(\mathbf{u})$  and  $\mathbf{u}$ , respectively. The proximal operator of a convex function  $\phi(\cdot) : \mathbb{C}^N \rightarrow \mathbb{R}$  is defined as

$$\text{prox}_{\phi}(\mathbf{u}) = \arg \min_{\mathbf{s} \in \mathbb{C}^N} \left\{ \phi(\mathbf{s}) + \frac{1}{2} \|\mathbf{s} - \mathbf{u}\|_2^2 \right\}. \quad (1)$$

Note that we use this definition formally even when  $\phi(\cdot)$  is a nonconvex function.

## II. SYSTEM MODEL

Here, we describe the system model considered in this paper. Fig. 1 shows the system model of the uplink communications in the IoT environments, which is modeled as MU-MIMO OFDM or SC-CP system. The number of antennas at the base station, the number of IoT terminals and the block length are denoted as  $M$ ,  $N$ , and  $Q$ , respectively. The frequency domain received signal model of the precoded MU-MIMO OFDM system can be written as

$$\begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_M \end{bmatrix} = \begin{bmatrix} \Lambda^{(1,1)} \mathbf{P} & \cdots & \Lambda^{(1,N)} \mathbf{P} \\ \vdots & & \vdots \\ \Lambda^{(M,1)} \mathbf{P} & \cdots & \Lambda^{(M,N)} \mathbf{P} \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_N \end{bmatrix} + \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_M \end{bmatrix}, \quad (2)$$

where  $\mathbf{r}_m \in \mathbb{C}^Q$  ( $m = 1, \dots, M$ ) is the received signal block in the frequency domain at the  $m$ -th antenna of the base station and  $\mathbf{s}_n \in \mathbb{C}^Q$  ( $n = 1, \dots, N$ ) is the transmitted signal block of the  $n$ -th IoT terminal.  $\mathbf{v}_m \in \mathbb{C}^Q$  is the additive white Gaussian noise in the frequency domain at the  $m$ -th antenna of the base station with mean of  $\mathbf{0}_Q$  and covariance matrix of  $\sigma_v^2 \mathbf{I}_Q$ . The elements of the frequency domain diagonal channel matrix  $\Lambda^{(m,n)} = \text{diag}(\lambda_1^{(m,n)}, \dots, \lambda_Q^{(m,n)}) \in \mathbb{C}^{Q \times Q}$  between the  $n$ -th IoT terminal and the  $m$ -th receiving antenna of the base station can be obtained by the discrete Fourier transform (DFT) of the  $L$ -path channel impulse response of  $\{h_1^{(m,n)}, \dots, h_L^{(m,n)}\}$  as

$$\begin{bmatrix} \lambda_1^{(m,n)} \\ \vdots \\ \lambda_Q^{(m,n)} \end{bmatrix} = \sqrt{Q} \mathbf{D} \begin{bmatrix} h_1^{(m,n)} \\ \vdots \\ h_L^{(m,n)} \\ \mathbf{0}_{Q-L} \end{bmatrix}, \quad (3)$$

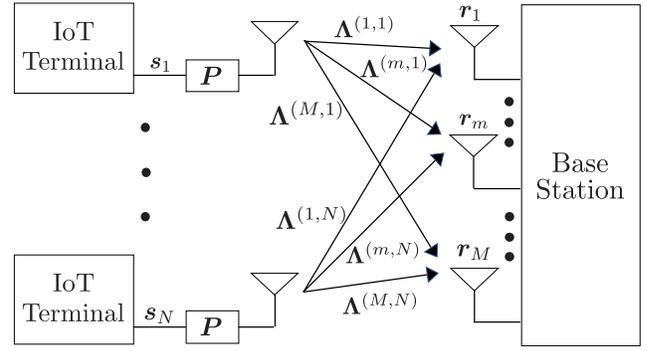


Fig. 1: Uplink MU-MIMO OFDM/SC-CP system model for base station and IoT terminals

where  $\mathbf{D}$  is a  $Q$ -point unitary DFT matrix.  $\mathbf{P} \in \mathbb{C}^{Q \times Q}$  represents the precoding matrix, which is common to all terminals. When  $\mathbf{P}$  has an arbitrary structure, (2) is a received signal model of precoded MU-MIMO OFDM. On the other hand, when we set  $\mathbf{P} = \mathbf{D}$ , it results in the received signal model of non-precoded MU-MIMO SC-CP [10], [15]. We assume that  $N_{\text{act}}$  IoT terminals out of  $N$  terminals are active. While  $N - N_{\text{act}}$  nonactive terminals actually keep silent, we can regard they transmit all zero signal block of  $\mathbf{0}_Q$ .

We define  $\mathbf{r} = [\mathbf{r}_1^T \cdots \mathbf{r}_M^T]^T \in \mathbb{C}^{QM}$  and  $\mathbf{s} = [\mathbf{s}_1^T \cdots \mathbf{s}_N^T]^T \in \mathbb{C}^{QN}$ , then (2) can be rewritten in a simpler form as

$$\mathbf{r} = \mathbf{A} \mathbf{s} + \mathbf{v}, \quad (4)$$

where  $\mathbf{A} \in \mathbb{C}^{QM \times QN}$  is the whole channel matrix defined as

$$\mathbf{A} = \begin{bmatrix} \Lambda^{(1,1)} \mathbf{P} & \cdots & \Lambda^{(1,N)} \mathbf{P} \\ \vdots & & \vdots \\ \Lambda^{(M,1)} \mathbf{P} & \cdots & \Lambda^{(M,N)} \mathbf{P} \end{bmatrix}, \quad (5)$$

and  $\mathbf{v} = [\mathbf{v}_1^T \cdots \mathbf{v}_M^T]^T \in \mathbb{C}^{QM}$ .

## III. IWSCSR-GS OPTIMIZATION PROBLEM

In this section, we present SCSR-GS optimization problem, which is a complex discrete-valued vector reconstruction method, and IWSCSR-GS optimization [11], which can improve the detection performance by iteratively adjusting the weights in the optimization problem.

Let  $\mathcal{S} = \{c_1, \dots, c_S\}$  be alphabet of the unknown signal. If we employ QPSK, we have  $S = 5$  including the symbol of 0 and  $\{c_1, \dots, c_5\} = \{0, 1 + j, -1 + j, 1 - j, -1 - j\}$ . The probability distribution of the elements of the unknown vector  $\mathbf{x} \in \mathbb{C}^{QN}$  is defined by

$$\Pr(x_i = c_\ell) = p_\ell \quad (\ell = 1, \dots, S, i = 1, \dots, QN), \quad (6)$$

where  $\sum_{\ell=1}^S p_\ell = 1$ . The SCSR-GS optimization problem [10], [11] is given by

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{C}^{QN}} \sum_{\ell=1}^S q_\ell g_\ell(\mathbf{x} - c_\ell \mathbf{1}_{QN}) + \alpha \sum_{n=1}^N \|\mathbf{x}_n\|_2 \\ \text{subject to } \|\mathbf{r} - \mathbf{A} \mathbf{x}\|_2 \leq \epsilon, \end{aligned} \quad (7)$$

where  $\alpha, \epsilon > 0$ , and  $q_\ell \geq 0$  ( $\ell = 1, \dots, S$ ) satisfies  $\sum_{\ell=1}^S q_\ell = 1$ . The vector  $\mathbf{x}_n \in \mathbb{C}^Q$  is the  $n$ -th subvector of  $\mathbf{x} = [\mathbf{x}_1^T \cdots \mathbf{x}_N^T] \in \mathbb{C}^{QN}$ . The function  $g_\ell(\cdot)$  is a sparse regularization function, where we assume that its proximal operator is easy to compute. In [10] and [11], the  $\ell_p$  norm ( $p = 0, 1, 1/2, 2/3$ ) or  $\ell_1 - \ell_2$  difference based functions have been employed as the function  $g_\ell(\cdot)$ . The first term in the optimization problem (7) can be considered as a discrete regularizer for  $\mathbf{x} \in \mathbb{C}^{QN}$ , which uses the fact that  $\mathbf{x} - c_\ell \mathbf{1}_{QN}$  is a sparse vector. The second term promotes the group sparsity of the solution. We adopt the  $\ell_2$  norm for this term, which is defined as

$$\|\mathbf{x}_n\|_2 = \sqrt{\sum_{i=1}^Q |x_{n,i}|^2} \quad (8)$$

for a complex vector  $\mathbf{x}_n = [x_{n,1} \cdots x_{n,Q}]^T \in \mathbb{C}^Q$ .

In order to take the prior distribution of unknown vectors into consideration, we have introduced weights for each symbol  $x_i$  in the SCSR-GS, and the resultant optimization problem is given by

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{C}^{QN}} \sum_{\ell=1}^S \sum_{i=1}^{QN} q_{i,\ell} g_\ell(x_i - c_\ell) + \sum_{n=1}^N \alpha_n \|\mathbf{x}_n\|_2 \\ \text{subject to } \|\mathbf{r} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon, \end{aligned} \quad (9)$$

where  $q_{i,\ell} \geq 0$  and  $\alpha_n > 0$  satisfy  $\sum_{\ell=1}^S q_{i,\ell} = 1$  for any  $i = 1, \dots, QN$  and  $\sum_{n=1}^N \alpha_n = N\alpha$ , respectively. The weights  $q_{i,\ell}$  and  $\alpha_n$  are determined according to the prior distribution of each symbol. It is true that we usually have no prior information on the distribution, and thus we have to use uniform distribution for the weights, which results in no performance gain against SCSR-GS. However, if we employ an iterative approach, we can obtain the prior distribution from the tentative estimate of the unknown vector in the previous iteration as in the case with the turbo signal processing, and use the distribution to update the weights in the optimization problem (9), which leads to performance improvement. We call this iterative solution for the optimization problem of (9) as IWSCSR-GS [11].

#### IV. PROPOSED METHOD

In this section, we propose IWSCSR-GS with SCAD or MCP, which are piecewise continuous nonconvex sparse regularizers. Specifically, we consider to use SCAD or MCP for  $g_\ell(\cdot)$  in the first discrete regularization term of (9) to improve the reconstruction performance of IWSCSR-GS.

Let a real vector be  $\mathbf{t} \in \mathbb{R}^N$ , then, SCAD [16] is defined as

$$f^{\text{SCAD}}(t_n; \gamma, a) = \begin{cases} \gamma |t_n| & (|t_n| \leq \gamma) \\ -\frac{t_n^2 - 2a\gamma|t_n| + \gamma^2}{2(a-1)} & (\gamma < |t_n| \leq a\gamma) \\ \frac{(a+1)\gamma^2}{2} & (|t_n| > a\gamma), \end{cases} \quad (10)$$

and MCP [17] is defined as

$$f^{\text{MCP}}(t_n; \gamma, a) = \begin{cases} \gamma |t_n| - \frac{t_n^2}{2a} & (|t_n| \leq a\gamma) \\ \frac{a\gamma^2}{2} & (|t_n| > a\gamma), \end{cases} \quad (11)$$

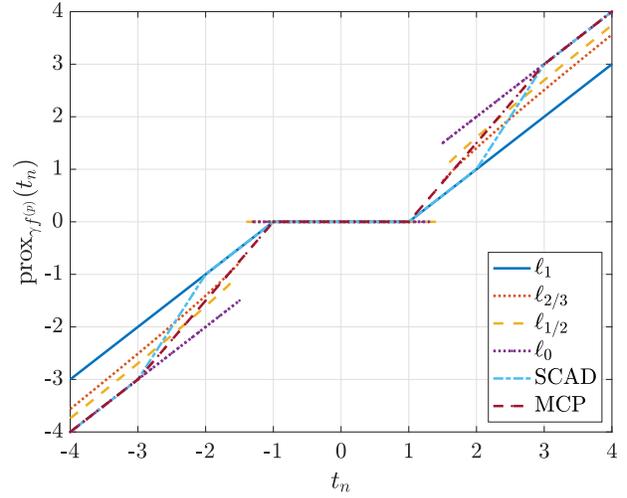


Fig. 2: Proximal operators for nonconvex regularizers ( $\gamma = 1, a = 3$ )

where  $\gamma \in (0, \infty)$  and  $a \in (1, \infty)$  are the nonconvexity parameters, which determine the properties of SCAD and MCP. It should be noted that both SCAD and MCP include  $\ell_1$  norm regularization as a special case where  $a$  of (10) and (11) goes to infinity. From (1), proximal operators for SCAD and MCP can be respectively calculated as

$$\begin{aligned} [\text{prox}_{\gamma f^{\text{SCAD}}}(\mathbf{t})]_n &= \begin{cases} t_n - \text{sgn}(t_n)\gamma & \text{for } 2\gamma \geq |t_n| > \gamma \\ \frac{a-1}{a-2}(t_n - \text{sgn}(t_n)\frac{a\gamma}{a-1}) & \text{for } a\gamma \geq |t_n| > 2\gamma \\ t_n & \text{for } |t_n| > a\gamma \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (12)$$

and

$$\begin{aligned} [\text{prox}_{\gamma f^{\text{MCP}}}(\mathbf{t})]_n &= \begin{cases} \frac{a}{a-1}(t_n - \text{sgn}(t_n)\gamma) & \text{for } a\gamma \geq |t_n| > \gamma \\ t_n & \text{for } |t_n| > a\gamma \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (13)$$

where  $\text{sgn}(t_n)$  is the sign of  $t_n$ . Fig. 2 shows examples of the proximal operators for SCAD, MCP and  $\ell_p$  norm, which is studied in [18]–[20]. The figure shows that the proximal operators for  $\ell_p$  norm functions ( $p = 0, 1/2, 2/3$ ) are discontinuous, while those for SCAD and MCP are continuous.

In accordance with [9] and [21], we can employ two forms of  $g_\ell(\cdot)$  of (9) given by

$$g_\star^{(\cdot)}(\mathbf{u}) = f^{(\cdot)}(|\mathbf{u}|), \quad (14)$$

$$g_{\star\star}^{(\cdot)}(\mathbf{u}) = f^{(\cdot)}(\text{Re}(\mathbf{u})) + f^{(\cdot)}(\text{Im}(\mathbf{u})), \quad (15)$$

where  $\mathbf{u} = [u_1, \dots, u_N]^T \in \mathbb{C}^N$  and  $f^{(\cdot)}$  denotes the regularization function in real space. As a result of the investigations in [9], [21], we have found that  $g_\star^{(\cdot)}$  should

**Algorithm 1** IWSCSR-GS

**Input:**  $\mathbf{r} \in \mathbb{C}^{QM}$ ,  $\mathbf{A} \in \mathbb{C}^{QM \times QN}$ 
**Output:**  $\hat{\mathbf{s}} \in \mathbb{C}^{QN}$ 


---

```

1: Initialize  $q_{i,\ell} = q_\ell$  ( $i = 1, \dots, QN$ ,  $\ell = 1, \dots, S$ ),  $\alpha_n = \alpha$  ( $n = 1, \dots, N$ )
2: for  $t = 1$  to  $T$  do
3:   Fix  $\alpha, \rho > 0$ ,  $\mathbf{z}_1^0, \dots, \mathbf{z}_S^0$ ,  $\mathbf{z}_{GS}^0$ ,  $\mathbf{w}_1^0, \dots, \mathbf{w}_S^0$ ,  $\mathbf{w}_{GS}^0 \in \mathbb{C}^{QN}$ ,  $\mathbf{z}_B^0$ ,  $\mathbf{w}_B^0 \in \mathbb{C}^{QM}$ 
4:   for  $k = 0$  to  $K - 1$  do
5:      $\mathbf{x}^{k+1} = ((S + 1)\mathbf{I}_{QN} + \mathbf{A}^H \mathbf{A})^{-1}$ 
6:        $\cdot (\sum_{\ell=1}^S (\mathbf{z}_\ell^k - \mathbf{w}_\ell^k) + (\mathbf{z}_{GS}^k - \mathbf{w}_{GS}^k)$ 
7:          $+ \mathbf{A}^H (\mathbf{z}_B^k - \mathbf{w}_B^k))$ 
8:      $\mathbf{y}_\ell^{k+1} = \mathbf{x}^{k+1} + \mathbf{w}_\ell^k$  ( $\ell = 1, \dots, S$ )
9:      $\mathbf{y}_{GS}^{k+1} = \mathbf{x}^{k+1} + \mathbf{w}_{GS}^k$ 
10:     $\mathbf{y}_B^{k+1} = \mathbf{A} \mathbf{x}^{k+1} + \mathbf{w}_B^k$ 
11:     $z_{i,\ell}^{k+1} = c_\ell + \text{prox}_{\frac{q_{i,\ell}}{2\rho} g_\ell}(\mathbf{y}_{i,\ell}^{k+1} - c_\ell)$  ( $i = 1, \dots, QN$ ,  $\ell = 1, \dots, S$ )
12:     $\mathbf{z}_{GS,n}^{k+1} = [\text{prox}_{\frac{q_n}{2\rho} g_{GS,n}}(\mathbf{y}_{GS}^{k+1})]_n$  ( $n = 1, \dots, N$ )
13:     $\mathbf{z}_B^{k+1} = \text{prox}_{\frac{1}{2\rho} \chi_B}(\mathbf{y}_B^{k+1})$ 
14:     $\mathbf{w}_\ell^{k+1} = \mathbf{y}_\ell^{k+1} - \mathbf{z}_\ell^{k+1}$  ( $\ell = 1, \dots, S$ )
15:     $\mathbf{w}_{GS}^{k+1} = \mathbf{y}_{GS}^{k+1} - \mathbf{z}_{GS}^{k+1}$ 
16:     $\mathbf{w}_B^{k+1} = \mathbf{y}_B^{k+1} - \mathbf{z}_B^{k+1}$ 
17:   end for
18:    $q_{i,\ell} = \frac{|x_i^K - c_\ell|^{-1}}{\sum_{\ell'=1}^S |x_i^K - c_{\ell'}|^{-1}}$  ( $i = 1, \dots, QN$ ,  $\ell = 1, \dots, S$ )
19:    $\alpha_n = \frac{\|\mathbf{x}_n^K\|_2^{-1}}{\sum_{n'=1}^N \|\mathbf{x}_{n'}^K\|_2^{-1}} N \alpha$  ( $n = 1, \dots, N$ )
20: end for
21:  $\hat{\mathbf{s}} = \mathbf{x}^K$ 

```

---

be used corresponding to symbol 0 and  $g_{**}^{(\cdot)}$  should be used corresponding to other symbols due to the fact that the real and imaginary parts of the symbol 0 are 0 at the same time, and this paper also follows. The proximal operators in complex space can be written in terms of the corresponding proximal operator in real space as

$$\left[ \text{prox}_{\gamma g_{*}^{(\cdot)}}(\mathbf{u}) \right]_n = \left[ \text{prox}_{\gamma f^{(\cdot)}}(|\mathbf{u}|) \right]_n \frac{u_n}{|u_n|}, \quad (16)$$

$$\left[ \text{prox}_{\gamma g_{**}^{(\cdot)}}(\mathbf{u}) \right]_n = \left[ \text{prox}_{\gamma f^{(\cdot)}}(\text{Re}(\mathbf{u})) \right]_n + j \left[ \text{prox}_{\gamma f^{(\cdot)}}(\text{Im}(\mathbf{u})) \right]_n. \quad (17)$$

Using these proximity operators, the ADMM based algorithms to solve SCSR-GS optimization and IWSCSR-GS optimization are shown in Algorithm 1, where  $\rho > 0$ ,  $K$  is the number of iterations of the algorithm and  $T$  is the total number of iterations. Note that, when  $T = 1$ , IWSCSR-GS is the same as SCSR-GS. The function  $g_{GS,n}(\cdot)$  in the algorithms is the regularization function corresponding to the second term of (9), which accounts for group sparsity, and we employ  $g_{GS,n}(\mathbf{z}) = \|\mathbf{z}_n\|_2$ . The indicator function  $\chi_B(\mathbf{z}) = \chi_B(\mathbf{A}\mathbf{x})$

corresponds to the constraint in (9) and is given by

$$\chi_B(\mathbf{z}) = \begin{cases} 0 & \text{if } \mathbf{z} \in \mathcal{B} \\ \infty & \text{if } \mathbf{z} \notin \mathcal{B}, \end{cases} \quad (18)$$

where  $\mathcal{B} = \{\mathbf{u} \in \mathbb{C}^{QN} : \|\mathbf{r} - \mathbf{u}\|_2 \leq \epsilon\}$ , in accordance with [11]. In Algorithm 1, we employ the weight update equations as in [10],

$$q_{i,\ell} = \frac{|\hat{x}_i - c_\ell|^{-1}}{\sum_{\ell'=1}^S |\hat{x}_i - c_{\ell'}|^{-1}} \quad (19)$$

and

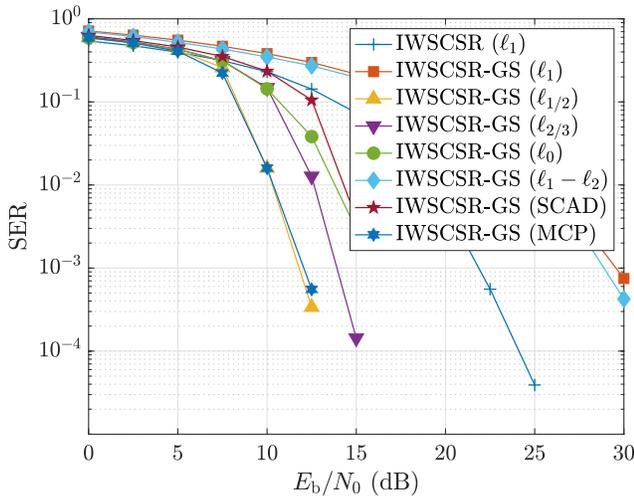
$$\alpha_n = \frac{\|\hat{\mathbf{x}}_n\|_2^{-1}}{\sum_{n'=1}^N \|\hat{\mathbf{x}}_{n'}\|_2^{-1}} N \alpha, \quad (20)$$

where we define  $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_N]^T = [\hat{x}_1, \dots, \hat{x}_{QN}]^T \in \mathbb{C}^{QN}$  as the estimate of  $\mathbf{s}$  at the previous iteration.

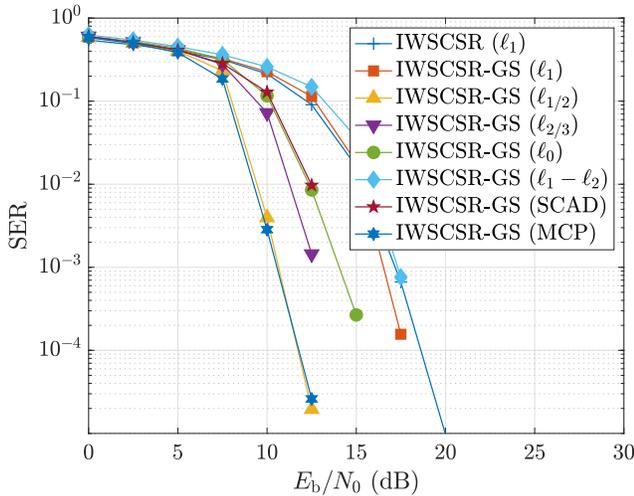
## V. NUMERICAL RESULTS

The signal detection performance of the proposed method is evaluated by the SER via computer simulations. The modulation scheme is QPSK, and the communication channels are assumed to be 10-path frequency selective channels. The block length is set to be  $Q = 64$  and all delayed signals are assumed to be within the cyclic prefix. Hadamard matrix is used for the MU-MIMO OFDM precoding matrix  $\mathbf{P}$ . In order to evaluate the performance of different system size, we have set the number of antennas at the base station  $M$ , the number of IoT terminals  $N$ , and the number of active IoT terminals  $N_{\text{act}}$  to be  $(M, N, N_{\text{act}}) = (6, 8, 7)$  for a small system and  $(M, N, N_{\text{act}}) = (60, 80, 70)$  for a large system, which correspond to the overloaded factor of  $N/M = 1.33$  in both cases. In these cases, the initial values of the weights are set to  $(q_1, q_2, q_3, q_4, q_5) = (\frac{N - N_{\text{act}}}{N}, \frac{N_{\text{act}}}{4N}, \frac{N_{\text{act}}}{4N}, \frac{N_{\text{act}}}{4N}, \frac{N_{\text{act}}}{4N}) = (1/8, 7/32, 7/32, 7/32, 7/32)$ . The parameter  $\rho$ , the weight  $\alpha$ , and the nonconvexity parameter  $a$  are set to be the best values for each methods in the trials. For comparison purpose, we have also evaluated the performance of IWSCSR-GS with  $\ell_p$  norm ( $p = 0, 1, 1/2, 2/3$ ) or  $\ell_1 - \ell_2$  difference [22] and the performance of iterative weighted SCSR (IWSCSR) [9] with  $\ell_1$  norm, which is a naive reconstruction method without using group sparsity, in the simulations.

Figs. 3–6 show the SER performance of IWSCSR-GS versus  $E_b/N_0$  (energy per bit to noise power spectral density ratio) for MU-MIMO OFDM. Figs. 3 and 4 show the results with the small system for  $T = 1$  and 5 in the proposed algorithm, respectively, and Figs. 5 and 6 show the results with the large system for  $T = 1$  and 5, respectively. In the case of the small system, the SER performance of the proposed IWSCSR-GS with SCAD is worse than or comparable to that of IWSCSR-GS with  $\ell_0$  norm, however in the case of the large system, the SER performance of IWSCSR-GS with SCAD exceeds that of IWSCSR-GS with  $\ell_0$  norm at  $E_b/N_0 = 15$  dB and that of IWSCSR-GS with SCAD achieves the best at  $E_b/N_0 = 15$  dB for  $T = 5$ . On the other hand, the proposed IWSCSR-GS with MCP has comparable performance to IWSCSR-GS with  $\ell_{1/2}$  norm, which has the



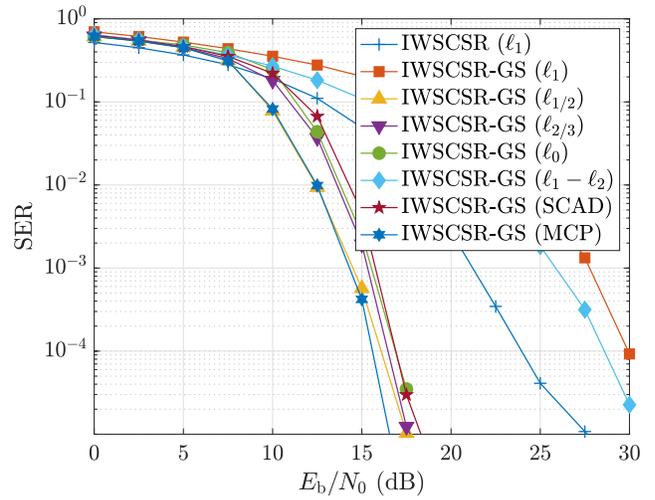
**Fig. 3:** SER performance (OFDM with precoding,  $M = 6, N = 8, N_{\text{act}} = 7, T = 1$ )



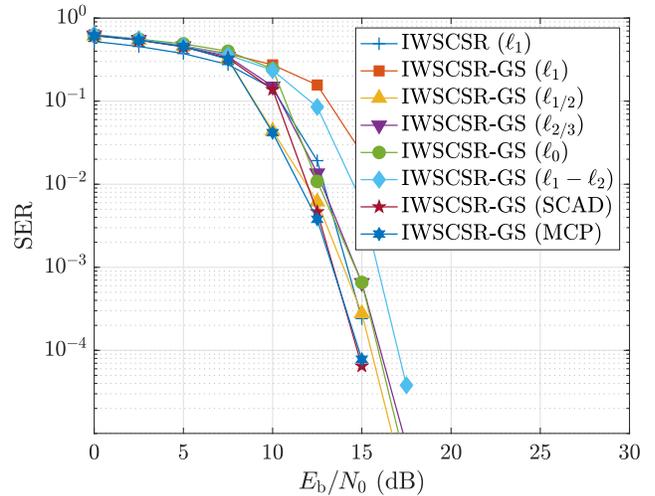
**Fig. 4:** SER performance (OFDM with precoding,  $M = 6, N = 8, N_{\text{act}} = 7, T = 5$ )

best performance among the existing methods, in the case of the small system, and furthermore, it achieves the best performance in the case of the large system except for the result of IWSCSR-GS with SCAD at  $E_b/N_0 = 15$  dB for  $T = 5$ .

Figs. 7–10 show the SER performance of IWSCSR-GS versus  $E_b/N_0$  for MU-MIMO SC-CP. Figs. 7 and 8 show the results with the small system for  $T = 1$  and 5 in the proposed algorithm, respectively, and Figs. 9 and 10 show the results with the large system for  $T = 1$  and 5, respectively. As is the case of MU-MIMO OFDM, we can see that the proposed IWSCSR-GS with SCAD has worse performance than or comparable performance to IWSCSR-GS with  $\ell_0$  norm for the small system and it achieves the best performance for the large system at  $E_b/N_0 = 15$  dB for  $T = 5$ . Similarly, the proposed IWSCSR-GS with MCP has comparable performance to IWSCSR-GS with  $\ell_{1/2}$  norm for



**Fig. 5:** SER performance (OFDM with precoding,  $M = 60, N = 80, N_{\text{act}} = 70, T = 1$ )



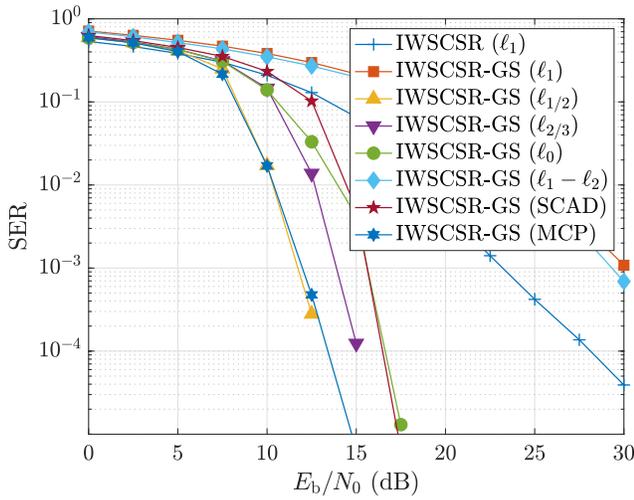
**Fig. 6:** SER performance (OFDM with precoding,  $M = 60, N = 80, N_{\text{act}} = 70, T = 5$ )

the small system, and it achieves the best performance for the large system except for the performance of IWSCSR-GS with SCAD at  $E_b/N_0 = 15$  dB for  $T = 5$ .

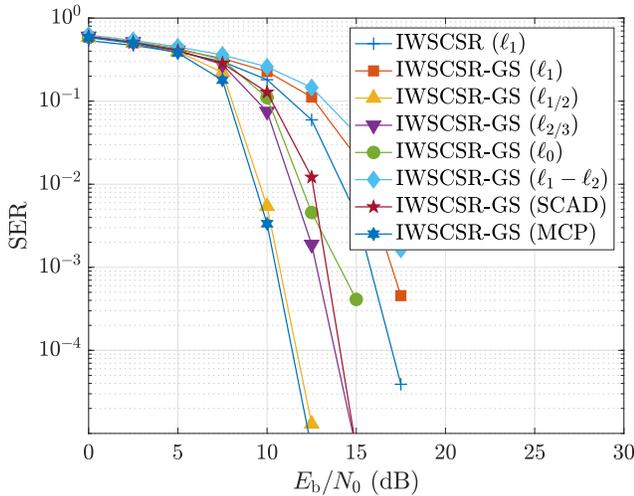
From Figs. 3–10, we can see that as the discrete regularizer in IWSCSR-GS, SCAD is suitable for large systems with high  $E_b/N_0$  regions, and MCP is the best choice for both small and large systems.

## VI. CONCLUSIONS

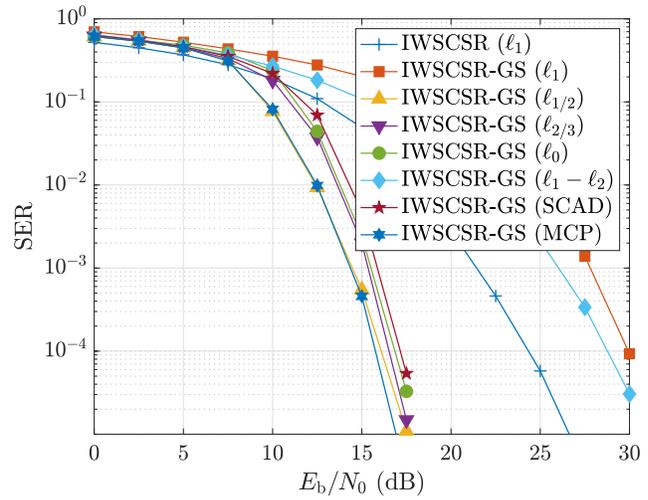
In this paper, we have considered an uplink overloaded MU-MIMO OFDM/SC-CP signal detection method using piecewise continuous nonconvex sparse regularizers, such as SCAD and MCP. Computer simulation results show that the proposed IWSCSR-GS with MCP achieves better SER performance than IWSCSR-GS with  $\ell_p$  norm ( $p = 0, 1, 1/2, 2/3$ ) or  $\ell_1 - \ell_2$  difference and that with SCAD achieves the best performance for large systems with high  $E_b/N_0$  region. Future



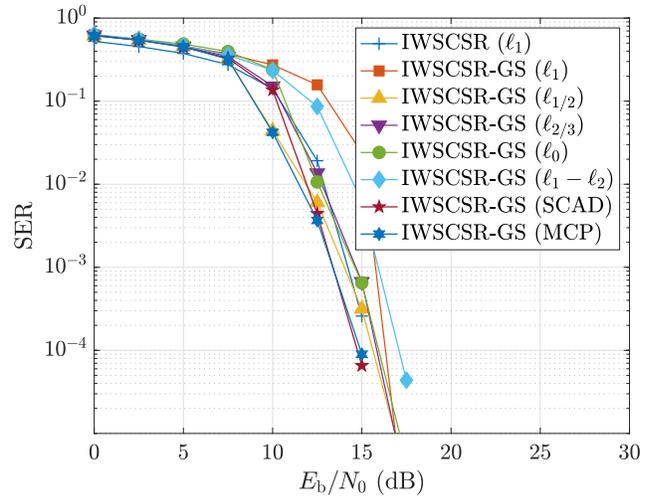
**Fig. 7:** SER performance (SC-CP without precoding,  $M = 6, N = 8, N_{\text{act}} = 7, T = 1$ )



**Fig. 8:** SER performance (SC-CP without precoding,  $M = 6, N = 8, N_{\text{act}} = 7, T = 5$ )



**Fig. 9:** SER performance (SC-CP without precoding,  $M = 60, N = 80, N_{\text{act}} = 70, T = 1$ )



**Fig. 10:** SER performance (SC-CP without precoding,  $M = 60, N = 80, N_{\text{act}} = 70, T = 5$ )

work includes detailed optimization of the parameters in the proposed algorithm and further investigation of the iterative signal detection including decoding of error correcting codes.

REFERENCES

- [1] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, July 2015.
- [2] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [3] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: applications, trends, technologies, and open research problems," *IEEE Netw.*, vol. 34, no. 3, pp. 134–142, May–June 2019.
- [4] Y. Nin, H. Matsuoka, Y. Sanada, "Performance comparison of overloaded MIMO system with and without antenna selection," *IEICE Trans. Commun.*, vol. E100-B, no. 5, pp. 762–770, May 2017.
- [5] K. Hayashi, A. Nakai-Kasai, R. Hayakawa, and S. Ha, "Uplink overloaded MU-MIMO OFDM signal detection methods using convex optimization," *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2018)*, pp. 1421–1427, Honolulu, USA, November 2018.
- [6] R. Hayakawa, K. Hayashi, "Convex optimization-based signal detection for massive overloaded MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7080–7091, November 2017.
- [7] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.
- [8] K. Hayashi, M. Nagahara, and T. Tanaka "A user's guide to compressed sensing for communications systems," *IEICE Trans. Commun.*, vol. E96-B, no. 3, pp. 685–712, March 2013.
- [9] R. Hayakawa and K. Hayashi, "Reconstruction of complex discrete-valued vector via convex optimization with sparse regularizers," *IEEE Access*, vol. 6, pp. 66 499–66 512, December 2018.
- [10] R. Hayakawa, A. Nakai-Kasai, and K. Hayashi, "Discreteness and Group Aparsity Aware Detection for Uplink Overloaded MU-MIMO Systems," *APSIPA Trans. Signal Inf. Process.*, vol. 9, no. E21, pp. 1-12, October 2020.
- [11] A. Nakai-Kasai, A. Hirayama, H. Honda, T. Sasaki, H. Yasukawa, R. Hayakawa, and K. Hayashi, "An Overloaded MU-MIMO OFDM/SC-CP Signal Detection Method using Non-Convex Sparse Regularizers and Group Sparsity," *IEICE Technical Report*, vol. 120, no. 292, RCC2020-11, pp. 1–6, December 2020 (in Japanese)
- [12] A. Sakata and Y. Xu, "Approximate message passing for nonconvex sparse regularization with stability and asymptotic analysis," *J. Stat.*

- Mech: Theory and Experiment*, vol. 2018, no. 3, 033404, March 2018.
- [13] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, January 1976.
  - [14] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, January 2011.
  - [15] K. Hayashi, A. Nakai-Kasai, A. Hirayama, H. Honda, T. Sasaki, H. Yasukawa, and R. Hayakawa, "An overloaded IoT signal detection method using non-convex sparse regularizers," Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2020), virtual conference, December 2020.
  - [16] J. Fan and R. Li, "Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties," *J. Amer. Statist. Assoc.*, vol. 96, no. 456, pp. 1348–1360, 2001.
  - [17] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Ann. Statist.*, vol. 38, no. 2, pp. 894–942, 2010.
  - [18] Z. Xu, X. Chang, F. Xu, and H. Zhang, " $L_{1/2}$  regularization: A thresholding representation theory and a fast solver," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1013–1027, July 2012.
  - [19] Y. Zhang and W. Ye, " $L_{2/3}$  regularization: Convergence of iterative thresholding algorithm," *J. Vis. Commun. Image Represent.*, vol. 33, pp. 350–357, 2015.
  - [20] F. Chen, L. Shen, and B. W. Suter, "Computing the proximity operator of the  $\ell_p$  norm with  $0 < p < 1$ " *IET Signal Process.*, vol. 10, no. 5, pp. 557–565, June 2016.
  - [21] R. Hayakawa and K. Hayashi, "Discrete-valued vector reconstruction by optimization with sum of sparse regularizers," in *Proc. European Signal Processing Conference (EUSIPCO)*, A Coruna, Spain, September 2019.
  - [22] Y. Lou and M. Yan, "Fast L1–L2 minimization via a proximal operator," *J. Sci. Comput.*, vol. 74, no. 2, pp. 767–785, February 2018.