ORIGINAL PAPER

Discreteness and group sparsity aware detection for uplink overloaded MU-MIMO systems

RYO HAYAKAWA, 🆔 AYANO NAKAI-KASAI 🆔 AND KAZUNORI HAYASHI 🆔

This paper proposes signal detection methods for frequency domain equalization (FDE) based overloaded multiuser multiple input multiple output (MU-MIMO) systems for uplink Internet of things (IoT) environments, where a lot of IoT terminals are served by a base station having less number of antennas than that of IoT terminals. By using the fact that the transmitted signal vector has the discreteness and the group sparsity, we propose a convex discreteness and group sparsity aware (DGS) optimization problem for the signal detection. We provide an optimization algorithm for the DGS optimization on the basis of the alternating direction method of multipliers (ADMM). Moreover, we extend the DGS optimization into weighted DGS (W-DGS) optimization and propose an iterative approach named iterative weighted DGS (IW-DGS), where we iteratively solve the W-DGS optimization problem with the update of the parameters in the objective function. We also discuss the computational complexity of the proposed IW-DGS and show that we can reduce the order of the complexity by using the structure of the channel matrix. Simulation results show that the symbol error rate (SER) performance of the proposed method is close to that of the oracle zero forcing (ZF) method, which perfectly knows the activity of each IoT terminal.

I. INTRODUCTION

Typical Internet of things (IoT) data collection environment using a base station with a large number of antennas in the 5th generation mobile communications systems (5G) [1, 2] can be considered as a special case of the massive multiuser multiple input multiple output (MU-MIMO) communications system. One of fundamental differences between the conventional massive MU-MIMO [3] and the IoT data collection is that the number of IoT nodes, in other words, the number of transmit antennas is typically much greater than that of receiving antennas even when a large antenna array is employed at the base station. Combined use with multiple access schemes will enable us to serve a large number of IoT nodes, but it also introduces additional delay, which may not be acceptable in many IoT applications.

MIMO systems having greater number of transmit antennas (more precisely, transmit streams) than that of receiving antennas is called overloaded MIMO [4–6]. The signal detection problem in such scenario is very difficult

due to the underdetermined nature of the problem. However, if the transmit symbol takes discrete values on a finite set (alphabet) as in the case of digital communications, we can detect signals on the basis of maximum likelihood (ML) approach. Since the ML approach is not tractable in the case of massive overloaded MIMO detection, the signal detection method based on the local neighborhood search has been proposed [7]. Moreover, since the computational complexity of the method is still too high when we use more than tens of transmit antennas, we have proposed a signal detection scheme using sum of absolute values (SOAV) optimization [5], which is based on the idea of convex optimization and compressed sensing [8, 9]. This approach is especially effective when the number of antennas is large and the number of elements in the alphabet is small.

Taking advantage of the fact that the transmission rate from each IoT node is typically low, we have proposed a detection scheme for uplink orthogonal frequency division multiplexing (OFDM) IoT signals [10], where the overloaded MIMO signal reconstruction scheme in [5, 11] is employed. However, the method cannot take the dependency between the real and the imaginary parts of the complex transmitted symbol into consideration because it is based on the SOAV optimization in the real-valued domain. In fact, it cannot use the fact that the real and the imaginary parts of the transmitted signal become 0 simultaneously for non-active terminals. To tackle this problem, in [12], we have employed a sparse complex discrete-valued vector

¹Graduate School of Engineering Science, Osaka University, Osaka 560-8531, Japan
²Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
³The Center for Innovative Research and Education in Data Science, Kyoto University, Kyoto 606-8315, Japan

Corresponding author:
Ryo Hayakawa
Email: rhayakawa@sys.es.osaka-u.ac.jp

reconstruction method named sum of complex sparse regularizers (SCSR) optimization [13]. In [12], we have focused on the overloaded MU-MIMO using single carrier block transmission with cyclic prefix (SC-CP) without precoding, which is equivalent to the MU-MIMO OFDM with precoding by a common discrete Fourier transform (DFT) matrix. In MU-MIMO SC-CP systems, neither the operations of inverse DFT (IDFT) nor the precoding is required at the transmitter, which can greatly reduce the complexity of IoT nodes. Moreover, the SC-CP signals have lower peak-to-average power ratio (PAPR) than the OFDM signals.

In this paper, we propose signal detection methods for uplink overloaded MU-MIMO systems on the basis of the preliminary conference papers [10, 12]. As mentioned above, the methods in [10, 12] use the discreteness of the transmitted signal vector to detect signals. However, since only a small number of terminals are typically active in the IoT transmission [14], the transmitted signal vector composed by concatenating transmitted signal vectors from all IoT nodes has not only the discreteness but also the group sparsity, i.e. most subvectors corresponding to the non-active terminals have all zero elements. To utilize such group sparsity, we propose discreteness and group sparsity aware (DGS) optimization, which uses two regularizers to promote both the discreteness and the group sparsity of the transmitted signal vector. The DGS optimization is a convex optimization problem in the complex-valued domain and hence it can also utilize the dependency between the real and the imaginary parts of the transmitted signals. We provide an optimization algorithm for the DGS optimization on the basis of the alternating direction method of multipliers (ADMM) [15–19]. Moreover, we extend the DGS optimization into the weighted DGS (W-DGS) optimization so that we can use the prior information of the transmitted signal vector. We then propose an iterative approach called iterative weighted DGS (IW-DGS), where we iterate the W-DGS optimization and the parameter update in the objective function. We also discuss the computational complexity of the proposed IW-DGS. By using the structure of the channel matrix, we can reduce the order of the complexity. Simulation results show that the proposed IW-DGS can achieve good symbol error rate (SER) performance close to the oracle zero forcing (ZF) method, which perfectly knows the activity of each IoT terminal. We also show that the performance of the proposed method in MU-MIMO OFDM is significantly improved by using the precoding with the Hadamard matrix or the DFT matrix.

The additional contributions of this paper against the preliminary conference papers [10, 12] are summarized as follows.

- We newly propose the DGS optimization and the corresponding ADMM-based algorithm, which utilizes both the discreteness and the group sparsity of the transmitted signal vector. On the other hand, the method in the preliminary papers [10, 12] uses only the discreteness. In fact, the proposed DGS optimization can be considered as an extension of the SCSR optimization [13] used in [12]. Simulation results show that the performance is significantly improved by using the group sparsity.

- We also propose the iterative approach using the W-DGS optimization, which can achieve much better performance than the DGS optimization.

- We discuss the computational complexity of the proposed method, and show that we can reduce the order of the complexity by using the structure of the channel matrix.

- We demonstrate that the performance of the proposed method is comparable to the oracle ZF method, which perfectly knows the support of the transmitted signal vector, i.e. active IoT terminals, via computer simulations.

The rest of the paper is organized as follows. We describe the overloaded MU-MIMO OFDM system and the overloaded MU-MIMO SC-CP system in Section II. In Section III, we propose the overloaded signal detection methods. Section IV shows simulation results to demonstrate the performance of the proposed approach. Finally, Section V gives some conclusions.

We use the following notations in this paper. $\mathbb{R}$ is the set of all real numbers and $\mathbb{C}$ is the set of all complex numbers. We denote the real part and the imaginary part by $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$, respectively. The transpose and the Hermitian transpose are indicated by $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{\mathsf{H}}$, respectively. We represent the imaginary unit by j, an $N \times N$ identity matrix by $\boldsymbol{I}_N$, the $M \times N$ zero matrix by $\mathbf{o}_{M \times N}$, the $N \times 1$ vector whose elements are all 1 by $\mathbf{1}_N$, and the $N \times 1$ vector whose elements are all 0 by $\mathbf{o}_N$. For a vector $\boldsymbol{u} = [u_1 \ldots u_N]^{\mathsf{T}} \in \mathbb{C}^N$, we define the ℓ_1 and ℓ_2 norms of $\boldsymbol{u}$ as $\|\boldsymbol{u}\|_1 = \sum_{n=1}^{N} |u_n|$ and $\|\boldsymbol{u}\|_2 = \sqrt{\sum_{n=1}^{N} |u_n|^2}$, respectively. $\mathrm{diag}(u_1, \ldots, u_N) \in \mathbb{C}^{N \times N}$ denotes the diagonal matrix whose (n, n) element is u_n. We represent the sign function by $\mathrm{sign}(\cdot)$. For a convex function $\phi : \mathbb{C}^N \to \mathbb{R} \cup \{\infty\}$, the proximity operator of ϕ is defined as

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

II. SYSTEM MODEL

In this section, we describe the system model considered in this paper. In Table 1, we summarize the notation for the system model.

A) Precoded MU-MIMO OFDM system

We consider uplink communications of IoT environments, which is modeled as a precoded MU-MIMO OFDM system. Figure 1 shows the system model, where the number of transmit terminals is N, the number of receiving antennas at the base station is M, and the number of subcarriers is Q. Given that the number of transmit terminal is typically large in IoT environments, we focus on the overloaded scenario and assume $M < N$ in this paper. The symbol alphabet and the frequency domain transmitted OFDM symbol vector from the n-th transmit IoT terminal are denoted by $\mathcal{S}$ and

Table 1. Notation for the system model

Notation	Definition
N	number of transmit terminals
N_{act}	number of active transmit terminals
M	number of receiving antennas
Q	number of subcarriers
L	length of channel impulse response
$\mathcal{S}$	symbol alphabet
$\boldsymbol{s}_n$	frequency domain transmitted OFDM symbol vector from the n-th transmit IoT terminal
$\boldsymbol{r}_m^{\text{f,OFDM}}$	frequency domain received OFDM signal vector at the m-th receiving antenna
$\boldsymbol{r}_m^{\text{f,SC-CP}}$	frequency domain received SC-CP signal vector at the m-th receiving antenna
$\boldsymbol{r}_m^{\text{t,SC-CP}}$	time domain received SC-CP signal vector at the m-th receiving antenna
$\lambda_q^{(m,n)}$	frequency response between the n-th IoT terminal and the m-th receiving antenna
$h_\ell^{(m,n)}$	impulse response between the n-th IoT terminal and the m-th receiving antenna
$\boldsymbol{P}$	precoding matrix
$\boldsymbol{D}$	DFT matrix
$\boldsymbol{v}_m^{\text{f}}$	frequency domain additive white noise vector at the m-th receiving antenna
$\boldsymbol{v}_m^{\text{t}}$	time domain additive white noise vector at the m-th receiving antenna
σ_{v}^2	noise variance
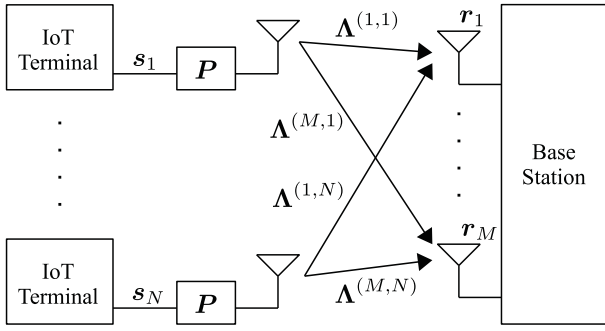
Fig. 1. Uplink MU-MIMO OFDM system for IoT environment.

$\boldsymbol{s}_n$, respectively. Here, taking IoT environment-specific feature into consideration, we assume only N_{act} IoT terminals out of N terminals are active meaning that only N_{act} terminals transmit OFDM signal blocks. Non-active $N - N_{\text{act}}$ terminals actually keep silent, but we can regard they transmit all zero signal block $\boldsymbol{o}_Q$. We thus have $\boldsymbol{s}_n \in \mathcal{S}^Q$ when the n-th terminal is active, and otherwise $\boldsymbol{s}_n = \boldsymbol{o}_Q$. When we use the cyclic prefix with the length greater than or equal to the channel order, the received signal vector after the removal of the cyclic prefix is given by

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

where $\boldsymbol{r}_m^{\text{f,OFDM}} \in \mathbb{C}^Q$ is the frequency domain received OFDM signal block at the m-th receiving antenna [10]. The diagonal matrix $\boldsymbol{\Lambda}^{(m,n)} = \text{diag}(\lambda_1^{(m,n)}, \ldots, \lambda_Q^{(m,n)}) \in \mathbb{C}^{Q \times Q}$ is composed of the channel frequency responses with the

order of $L - 1$ between the n-th IoT terminal and the m-th receiving antenna. The diagonal elements can be written as

$$
\begin{bmatrix} \lambda_1^{(m,n)} \\ \vdots \\ \lambda_Q^{(m,n)} \end{bmatrix} = \sqrt{Q}\boldsymbol{D} \begin{bmatrix} h_1^{(m,n)} \\ \vdots \\ h_L^{(m,n)} \\ \boldsymbol{o}_{Q-L} \end{bmatrix}, \quad (2)
$$

where $\boldsymbol{D} \in \mathbb{C}^{Q \times Q}$ is a Q-point unitary DFT matrix defined as

$$
\boldsymbol{D} = \frac{1}{\sqrt{Q}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-j\frac{2\pi \times 1 \times 1}{Q}} & \cdots & e^{-j\frac{2\pi \times 1 \times (Q-1)}{Q}} \\ \vdots & \vdots & & \vdots \\ 1 & e^{-j\frac{2\pi \times (Q-1) \times 1}{Q}} & \cdots & e^{-j\frac{2\pi \times (Q-1) \times (Q-1)}{Q}} \end{bmatrix} \quad (3)
$$

and $h_1^{(m,n)}, \ldots, h_L^{(m,n)}$ denotes the impulse response of the frequency-selective channel between the n-th IoT terminal and m-th receiving antenna. $\boldsymbol{P} \in \mathbb{C}^{Q \times Q}$ is a precoding matrix required to achieve good detection performance by the convex optimization-based detection scheme. Note that, in [10], we have numerically confirmed that we do not have to use different precoding matrices among IoT terminals, and a common Hadamard precoding matrix leads to good SER performance. $\boldsymbol{v}_m^{\text{f}} \in \mathbb{C}^Q$ is the frequency domain additive white noise vector at the m-th receiving antenna with mean $\boldsymbol{o}_Q$ and covariance matrix $\sigma_{\text{v}}^2 \boldsymbol{I}_Q$.

B) Non-precoded MU-MIMO SC-CP system

Here, we show a non-precoded MU-MIMO SC-CP signal model. Assuming the length of cyclic prefix is greater than or equal to the channel order $L - 1$, the time domain received signal block at the m-th receiving antenna of the base station is written as

$$
\boldsymbol{r}_m^{\text{t,SC-CP}} = \sum_{n=1}^{N} \boldsymbol{D}^{\mathsf{H}} \boldsymbol{\Lambda}^{(m,n)} \boldsymbol{D}\boldsymbol{s}_n + \boldsymbol{v}_m^{\text{t}}, \quad (4)
$$

where $\boldsymbol{v}_m^{\text{t}} \in \mathbb{C}^Q$ is the time domain additive white noise vector at the m-th receiving antenna having mean $\boldsymbol{o}_Q$ and covariance matrix $\sigma_{\text{v}}^2 \boldsymbol{I}_Q$ [12, 20]. By stacking from $\boldsymbol{r}_1^{\text{t,SC-CP}}$ to $\boldsymbol{r}_M^{\text{t,SC-CP}}$ in (4), and multiplying a unitary matrix of

$$
\begin{bmatrix} \boldsymbol{D} & \boldsymbol{o} & \cdots & \boldsymbol{o} \\ \boldsymbol{o} & \boldsymbol{D} & & \vdots \\ \vdots & & \ddots & \boldsymbol{o} \\ \boldsymbol{o} & \cdots & \boldsymbol{o} & \boldsymbol{D} \end{bmatrix} \in \mathbb{C}^{QM \times QM} \quad (5)
$$

from the left of both sides, we have the frequency domain received non-precoded SC-CP IoT signal vector at the base

station as

$$
\begin{bmatrix} r_1^{\mathrm{f,SC\text{-}CP}} \\ \vdots \\ r_M^{\mathrm{f,SC\text{-}CP}} \end{bmatrix} = \begin{bmatrix} D & o & \cdots & o \\ o & D & & \vdots \\ \vdots & & \ddots & o \\ o & \cdots & o & D \end{bmatrix} \begin{bmatrix} r_1^{\mathrm{t,SC\text{-}CP}} \\ \vdots \\ r_M^{\mathrm{t,SC\text{-}CP}} \end{bmatrix} \quad (6)
$$

$$
= \begin{bmatrix} \Lambda^{(1,1)}D & \cdots & \Lambda^{(1,N)}D \\ \vdots & & \vdots \\ \Lambda^{(M,1)}D & \cdots & \Lambda^{(M,N)}D \end{bmatrix} \begin{bmatrix} s_1 \\ \vdots \\ s_N \end{bmatrix} + \begin{bmatrix} v_1^{\mathrm{f}} \\ \vdots \\ v_M^{\mathrm{f}} \end{bmatrix}, \quad (7)
$$

where $r_m^{\mathrm{f,SC\text{-}CP}} \in \mathbb{C}^Q$ is the frequency domain received SC-CP signal vector at the m-th base station antenna and $v_m^{\mathrm{f}} = D v_m^{\mathrm{t}}$ $(m = 1, \ldots, M)$. It should be noted here that this received signal model can be regarded as a special case of (1), where the precoding matrix P is set to be D. Thus, if DFT matrix D is appropriate for the precoding matrix of overloaded MU-MIMO OFDM system with the convex optimization-based signal detection, then the choice of non-precoded SC-CP signaling is extremely suited for IoT environments because this approach requires neither the IDFT operation nor the precoding operation at the IoT node (transmitter side).

III. PROPOSED SIGNAL DETECTION METHOD

The MU-MIMO OFDM signal model (1) and the MU-MIMO SC-CP signal model (7) can be represented as

$$
r = As + v. \quad (8)
$$

Specifically, we have

$$
r = \left[\left(r_1^{\mathrm{f,OFDM}} \right)^{\mathsf{T}} \cdots \left(r_M^{\mathrm{f,OFDM}} \right)^{\mathsf{T}} \right]^{\mathsf{T}} \in \mathbb{C}^{QM} \quad (9)
$$

in MU-MIMO OFDM systems and

$$
r = \left[\left(r_1^{\mathrm{f,SC\text{-}CP}} \right)^{\mathsf{T}} \cdots \left(r_M^{\mathrm{f,SC\text{-}CP}} \right)^{\mathsf{T}} \right]^{\mathsf{T}} \in \mathbb{C}^{QM} \quad (10)
$$

in MU-MIMO SC-CP systems. A is the whole channel matrix given by

$$
A = \begin{bmatrix} \Lambda^{(1,1)}P & \cdots & \Lambda^{(1,N)}P \\ \vdots & & \vdots \\ \Lambda^{(M,1)}P & \cdots & \Lambda^{(M,N)}P \end{bmatrix}, \quad (11)
$$

where $P = D$ in the case of MU-MIMO SC-CP. $s = [s_1^{\mathsf{T}} \cdots s_N^{\mathsf{T}}]^{\mathsf{T}} \in \mathbb{C}^{QN}$ and $v = \left[\left(v_1^{\mathrm{f}} \right)^{\mathsf{T}} \cdots \left(v_M^{\mathrm{f}} \right)^{\mathsf{T}} \right]^{\mathsf{T}} \in \mathbb{C}^{QM}$ are the transmit symbol vector and the additive noise vector, respectively.

In this section, we describe the proposed signal detection method to estimate the transmitted symbol vector s from the received signal vector r and the channel matrix A in (8).

A) Discreteness and group sparsity of transmitted symbol vector

One of the main ideas of the proposed method is based on the fact that the transmitted symbol has the discreteness, i.e. the element of s is in the finite set $\mathcal{S} \cup \{o\}$. For the reconstruction of such discrete-valued vector, several methods have been proposed [21–24]. These methods reconstruct the discrete-valued vector in the real domain. For the reconstruction of the complex discrete-valued vector, SCSR optimization [13] has been proposed by extending the approach in [24]. For details of the related methods, see Section III-F.

When only a few transmit terminals are active, the transmitted symbol vector s has not only the discreteness but also the group sparsity, i.e. $s_n = o_Q$ holds for most transmit terminals, because the transmitted symbol vector of non-active terminals can be regarded as o_Q. Note that not only is s sparse, but multiple transmitted symbols from the non-active IoT terminal become zero simultaneously when OFDM or SC-CP signaling is employed. Such group sparsity has been utilized as prior knowledge in the literature of compressed sensing and sparse regression [25–27] as well as wireless communications [28, 29]. Thus, we can expect a certain improvement of the detection performance by using the group sparsity of the transmitted signal vector s.

B) Proposed DGS optimization

The proposed method utilizes both the discreteness and the group sparsity of s discussed in the previous subsection. The proposed DGS optimization problem is given by

$$
\underset{x \in \mathbb{C}^{QN}}{\text{minimize}} \left\{ \sum_{\ell=1}^{S} q_\ell g_\ell \left(x - c_\ell \mathbf{1} \right) + \alpha \sum_{n=1}^{N} \|x_n\|_2 \right\}
$$
$$
\text{subject to } \|r - Ax\|_2 \le \varepsilon, \quad (12)
$$

where $c_1, \ldots, c_S$ denotes all elements in $\{o\} \cup \mathcal{S}$. For example, when we use quadrature phase shift keying (QPSK) and set $\mathcal{S} = \{1+j, -1+j, -1-j, 1-j\}$, we have $S = 5$ and $c_1 = o, c_2 = 1+j, c_3 = -1+j, c_4 = -1-j, c_5 = 1-j$. The vector $x_n \in \mathbb{C}^Q$ is the n-th subvector of $x = \left[x_1^{\mathsf{T}} \cdots x_N^{\mathsf{T}} \right]^{\mathsf{T}} \in \mathbb{C}^{QN}$. q_ℓ $(> o)$, α $(> o)$, and ε $(> o)$ are parameters $(\ell = 1, \ldots, S)$, where $\sum_{\ell=1}^{S} q_\ell = 1$. The function $g_\ell(\cdot)$ is a sparse regularizer for a sparse vector in the complex-valued domain. In this paper, we consider two convex regularizers given by

$$
h_1(u) = \|u\|_1 \quad (13)
$$
$$
= \sum_{i=1}^{QN} \sqrt{\mathrm{Re}\{u_i\}^2 + \mathrm{Im}\{u_i\}^2}, \quad (14)
$$
$$
h_2(u) = \|\mathrm{Re}\{u\}\|_1 + \|\mathrm{Im}\{u\}\|_1 \quad (15)
$$
$$
= \sum_{i=1}^{QN} \left(|\mathrm{Re}\{u_i\}| + |\mathrm{Im}\{u_i\}| \right) \quad (16)
$$

in accordance with [13]. The first one $h_1(\cdot)$ is based on the modulus of the complex number, whereas the second one

$h_2(\cdot)$ considers the real part and the imaginary part separately. We can also use some non-convex sparse regularizers as in [30], though the global optimizer cannot be necessarily obtained.

The objective function of the DGS optimization (12) is the weighted sum of two regularizers. The first term $\sum_{\ell=1}^{S} q_\ell g_\ell(\boldsymbol{x} - c_\ell \boldsymbol{1})$ in (12) has been proposed in [13] as a regularizer for the discrete-valued vector. On the basis of the idea of compressed sensing [8, 9], the sparse regularizer $g_\ell(\cdot)$ is applied by using the fact that $\boldsymbol{s} - c_\ell \boldsymbol{1}$ has some zero elements. A good choice of the sparse regularizer depends on the symbol alphabet $\mathcal{S}$. For example, when we use QPSK and set $c_1 = 0, c_2 = 1 + j, c_3 = -1 + j, c_4 = -1 - j, c_5 = 1 - j$, it is reasonable to define $g_1(\cdot) = h_1(\cdot)$ and $g_\ell(\cdot) = h_2(\cdot)$ ($\ell = 2, \dots, 5$). For more detailed discussion, see [13]. In addition to the first regularizer, we also use the regularizer $\alpha \sum_{n=1}^{N} \|\boldsymbol{x}_n\|_2$ to promote the group sparsity of $\boldsymbol{s}$. Note that we use the ℓ_2 norm $\|\boldsymbol{x}_n\|_2 = \sqrt{\sum_{i=1}^{Q} |x_{n,i}|^2}$ of the complex-valued vector $\boldsymbol{x}_n = [x_{n,1} \; \cdots \; x_{n,Q}]^\mathsf{T} \in \mathbb{C}^Q$, whereas the ℓ_2 norm of the real valued vector is commonly used in the literature [25–27].

It should be noted that the conventional method [10] considers an optimization problem in the real domain, whereas the DGS optimization problem (12) is the optimization in the complex domain $\mathbb{C}^{QN}$. Such approach enables us to directly utilize the discrete nature of the transmitted symbols in the complex domain. Moreover, the DGS optimization can utilize the group sparsity of the transmitted symbol vector, which has not been used in the conventional methods.

C) Proposed algorithm for DGS optimization based on ADMM

We here propose an algorithm for the DGS optimization. Since the objective function of the DGS optimization problem (12) is not differentiable, simple gradient-based methods cannot be applied. However, we can derive an optimization algorithm for the DGS optimization (12) on the basis of ADMM [15–19] by appropriate reformulation as shown below. We firstly rewrite the DGS optimization problem (12) with new variables $\boldsymbol{z}_1, \dots, \boldsymbol{z}_S, \boldsymbol{z}_{GS} \in \mathbb{C}^{QN}$ and $\boldsymbol{z}_B \in \mathbb{C}^{QM}$ as

$$
\underset{\substack{\boldsymbol{x}, \boldsymbol{z}_1, \dots, \boldsymbol{z}_S, \boldsymbol{z}_{GS} \in \mathbb{C}^{QN} \\ \boldsymbol{z}_B \in \mathbb{C}^{QM}}}{\text{minimize}} \left\{ \sum_{\ell=1}^{S} q_\ell g_\ell(\boldsymbol{z}_\ell - c_\ell \boldsymbol{1}) \right.
$$
$$
\left. + \alpha \sum_{n=1}^{N} \|\boldsymbol{z}_{GS,n}\|_2 + \chi_\mathcal{B}(\boldsymbol{z}_B) \right\}
$$
$$
\text{subject to} \quad \boldsymbol{x} = \boldsymbol{z}_1 = \cdots = \boldsymbol{z}_L = \boldsymbol{z}_{GS},
$$
$$
\boldsymbol{A}\boldsymbol{x} = \boldsymbol{z}_B. \tag{17}
$$

In (17), we express the constraint $\|\boldsymbol{r} - \boldsymbol{A}\boldsymbol{x}\|_2 \le \varepsilon$ in (12) by $\chi_\mathcal{B}(\boldsymbol{z}_B) = \chi_\mathcal{B}(\boldsymbol{A}\boldsymbol{x})$ in the objective function, where

we define $\mathcal{B} := \{\boldsymbol{u} \in \mathbb{C}^{QM} \mid \|\boldsymbol{r} - \boldsymbol{u}\|_2 \le \varepsilon\}$ and the corresponding indicator function

$$
\chi_\mathcal{B}(\boldsymbol{z}_B) = \begin{cases} 0 & (\boldsymbol{z}_B \in \mathcal{B}) \\ \infty & (\boldsymbol{z}_B \notin \mathcal{B}) \end{cases}. \tag{18}
$$

Since the objective function in (17) becomes infinity when $\boldsymbol{x}$ does not satisfy the constraint $\|\boldsymbol{r} - \boldsymbol{A}\boldsymbol{x}\|_2 \le \varepsilon$, the optimization problem (17) is equivalent to the original optimization problem (12). The vector $\boldsymbol{z}_{GS,n} \in \mathbb{C}^Q$ is the n-th subvector of $\boldsymbol{z}_{GS} = [\boldsymbol{z}_{GS,1}^\mathsf{T} \; \cdots \; \boldsymbol{z}_{GS,N}^\mathsf{T}]^\mathsf{T} \in \mathbb{C}^{QN}$. To further rewrite the optimization problem (17), we denote the objective function as

$$
g_{GS}(\boldsymbol{z}_{GS}) = \sum_{n=1}^{N} \|\boldsymbol{z}_{GS,n}\|_2, \tag{19}
$$

$$
f(\boldsymbol{z}) = \sum_{\ell=1}^{S} q_\ell g_\ell(\boldsymbol{z}_\ell - c_\ell \boldsymbol{1}) + \alpha g_{GS}(\boldsymbol{z}_{GS}) + \chi_\mathcal{B}(\boldsymbol{z}_B). \tag{20}
$$

The constraint in (17) can also be summarized as $\boldsymbol{\Phi}\boldsymbol{x} = \boldsymbol{z}$, where we define

$$
\boldsymbol{z} = [\boldsymbol{z}_1^\mathsf{T} \; \cdots \; \boldsymbol{z}_S^\mathsf{T} \; \boldsymbol{z}_{GS}^\mathsf{T} \; \boldsymbol{z}_B^\mathsf{T}]^\mathsf{T}
$$
$$
\in \mathbb{C}^{(S+1)QN+QM} \tag{21}
$$

$$
\boldsymbol{\Phi} = [\boldsymbol{I}_{QN} \; \cdots \; \boldsymbol{I}_{QN} \; \boldsymbol{I}_{QN} \; \boldsymbol{A}^\mathsf{T}]^\mathsf{T}
$$
$$
\in \mathbb{C}^{((S+1)QN+QM) \times QN}. \tag{22}
$$

By using the above expression, the DGS optimization problem (17) can be finally represented as

$$
\underset{\substack{\boldsymbol{x} \in \mathbb{C}^{QN} \\ \boldsymbol{z} \in \mathbb{C}^{(S+1)QN+QM}}}{\text{minimize}} \quad f(\boldsymbol{z})
$$
$$
\text{subject to} \quad \boldsymbol{\Phi}\boldsymbol{x} = \boldsymbol{z}. \tag{23}
$$

The optimization problem (23) is a standard form for ADMM, and hence we can derive the optimization algorithm based on ADMM.

The update equations of ADMM for (23) are given by

$$
\boldsymbol{x}^{k+1} = \underset{\boldsymbol{x} \in \mathbb{C}^{QN}}{\arg\min} \left\{ \rho \|\boldsymbol{\Phi}\boldsymbol{x} - \boldsymbol{z}^k + \boldsymbol{w}^k\|_2^2 \right\}, \tag{24}
$$

$$
\boldsymbol{y}^{k+1} = \boldsymbol{\Phi}\boldsymbol{x}^{k+1} + \boldsymbol{w}^k, \tag{25}
$$

$$
\boldsymbol{z}^{k+1} = \underset{\boldsymbol{z} \in \mathbb{C}^{(S+1)QN+QM}}{\arg\min} \left\{ f(\boldsymbol{z}) + \rho \|\boldsymbol{y}^{k+1} - \boldsymbol{z}\|_2^2 \right\}, \tag{26}
$$

$$
\boldsymbol{w}^{k+1} = \boldsymbol{y}^{k+1} - \boldsymbol{z}^{k+1}, \tag{27}
$$

where $\rho\;(> 0)$ is a parameter and $\boldsymbol{w}^k \in \mathbb{C}^{(S+1)QN+QM}$. In the algorithm, $\boldsymbol{x}^k$ is the estimate for the transmitted signal vector $\boldsymbol{s}$ at the k-th iteration. The Wirtinger derivative [31] of

$\left\| \mathbf{\Phi} x - z^k + w^k \right\|_2^2$ in (24) with respect to x is given by

$$\frac{\partial}{\partial x^{\mathsf{H}}} \left\{ \left\| \mathbf{\Phi} x - z^k + w^k \right\|_2^2 \right\}$$
$$= \left((S+1) I_{QN} + A^{\mathsf{H}} A \right) x$$
$$- \left(\sum_{\ell=1}^{S} (z_\ell^k - w_\ell^k) + (z_{\mathrm{GS}}^k - w_{\mathrm{GS}}^k) + A^{\mathsf{H}} (z_{\mathrm{B}}^k - w_{\mathrm{B}}^k) \right). \tag{28}$$

Here, in the same manner as z in (21), we define $w_1^k, \ldots, w_S^k, w_{\mathrm{GS}}^k \in \mathbb{C}^{QN}$ and $w_{\mathrm{B}}^k \in \mathbb{C}^{QM}$ as the subvectors of the vector

$$w^k = \left[w_1^{k\mathsf{T}} \cdots w_S^{k\mathsf{T}} w_{\mathrm{GS}}^{k\mathsf{T}} w_{\mathrm{B}}^{k\mathsf{T}} \right]^{\mathsf{T}}$$
$$\in \mathbb{C}^{(S+1)QN+QM}. \tag{29}$$

From (28), $(\partial / \partial x^{\mathsf{H}}) \{ \| \mathbf{\Phi} x - z^k + w^k \|_2^2 \} = \mathbf{0}$ gives the explicit formula for the update of x^k in (24) as

$$x^{k+1} = \left((S+1) I_{QN} + A^{\mathsf{H}} A \right)^{-1} \left(\sum_{\ell=1}^{S} \left(z_\ell^k - w_\ell^k \right) \right.$$
$$\left. + \left(z_{\mathrm{GS}}^k - w_{\mathrm{GS}}^k \right) + A^{\mathsf{H}} \left(z_{\mathrm{B}}^k - w_{\mathrm{B}}^k \right) \right). \tag{30}$$

The update of y^k in (25) can be expressed separately as

$$y^{k+1} = \begin{bmatrix} y_1^k \\ \vdots \\ y_S^k \\ y_{\mathrm{GS}}^k \\ y_{\mathrm{B}}^k \end{bmatrix} = \begin{bmatrix} x + w_1^k \\ \vdots \\ x + w_S^k \\ x + w_{\mathrm{GS}}^k \\ A x + w_{\mathrm{B}}^k \end{bmatrix}. \tag{31}$$

Next, we consider the update of z^k in (26), which can be denoted as

$$z^{k+1} = \mathrm{prox}_{(1/2\rho)f} \left(y^{k+1} \right) \tag{32}$$

$$= \begin{bmatrix} c_1 \mathbf{1} + \mathrm{prox}_{(q_1/2\rho)g_1} \left(y_1^{k+1} - c_1 \mathbf{1} \right) \\ \vdots \\ c_S \mathbf{1} + \mathrm{prox}_{(q_S/2\rho)g_S} \left(y_L^{k+1} - c_L \mathbf{1} \right) \\ \mathrm{prox}_{(\alpha/2\rho)g_{\mathrm{GS}}} \left(y_{\mathrm{GS}}^{k+1} \right) \\ \mathrm{prox}_{(1/2\rho)\chi_{\mathcal{B}}} \left(y_{\mathrm{B}}^{k+1} \right) \end{bmatrix} \tag{33}$$

because the function $f(\cdot)$ is separable as in (20), where we have also used the property of proximity operator for translation [17]. As discussed in [13], the proximity operator of

$\gamma g_\ell(\cdot)$ can be computed by using

$$\left[\mathrm{prox}_{\gamma h_1}(u) \right]_i = \begin{cases} (|u_i| - \gamma) \dfrac{u_i}{|u_i|} & (|u_i| \geq \gamma) \\ 0 & (|u_i| < \gamma) \end{cases}, \tag{34}$$

$$\left[\mathrm{prox}_{\gamma h_2}(u) \right]_i = \mathrm{sign} \left(\mathrm{Re}\{u_i\} \right) \max \left(|\mathrm{Re}\{u_i\}| - \gamma, 0 \right)$$
$$+ j \cdot \mathrm{sign} \left(\mathrm{Im}\{u_i\} \right) \max \left(|\mathrm{Im}\{u_i\}| - \gamma, 0 \right), \tag{35}$$

where $\gamma > 0$ and $u = [u_1 \cdots u_{QN}]^{\mathsf{T}} \in \mathbb{C}^{QN}$. We denote the i-th element of the vector by $[\cdot]_i$. The proximity operator of $\gamma g_{\mathrm{GS}}(\cdot)$ is given by

$$\left[\mathrm{prox}_{\gamma g_{\mathrm{GS}}}(u) \right]_n = \begin{cases} (\|u_n\|_2 - \gamma) \dfrac{u_n}{\|u_n\|_2} & (\|u_n\|_2 \geq \gamma) \\ \mathbf{0}_Q & (\|u_n\|_2 < \gamma) \end{cases}, \tag{36}$$

where $u_n \in \mathbb{C}^Q$ and $[\mathrm{prox}_{\gamma g_{\mathrm{GS}}}(u)]_n \in \mathbb{C}^Q$ are the n-th subvectors of $u = \left[u_1^{\mathsf{T}} \cdots u_N^{\mathsf{T}} \right]^{\mathsf{T}} \in \mathbb{C}^{QN}$ and $\mathrm{prox}_{\gamma g_{\mathrm{GS}}}(u)$, respectively. The proximity operator of $\gamma \chi_{\mathcal{B}}(\cdot)$ can be written as the projection onto $\mathcal{B}$, i.e.

$$\mathrm{prox}_{\gamma \chi_{\mathcal{B}}}(u) = \begin{cases} r + \varepsilon \dfrac{u - r}{\|u - r\|_2} & (\|u - r\|_2 \geq \varepsilon) \\ u & (\|u - r\|_2 < \varepsilon) \end{cases}. \tag{37}$$

We summarize the proposed algorithm for the DGS optimization (23) in Algorithm 1.

Algorithm 1 Proposed Algorithm for DGS Optimization (23)

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

1: Fix $\rho > 0$, $z_1^0, \ldots, z_L^0, z_{\mathrm{GS}}^0, w_1^0, \ldots, w_L^0, w_{\mathrm{GS}}^0 \in \mathbb{C}^{QN}$, $z_{\mathrm{B}}^0, w_{\mathrm{B}}^0 \in \mathbb{C}^{QM}$
2: **for** $k = 0$ to $K_{\mathrm{itr}} - 1$ **do**
3: $\quad x^{k+1} = \left((S+1) I_{QN} + A^{\mathsf{H}} A \right)^{-1}$
4: $\quad\quad \times \left(\sum_{\ell=1}^S \left(z_\ell^k - w_\ell^k \right) + \left(z_{\mathrm{GS}}^k - w_{\mathrm{GS}}^k \right) \right.$
5: $\quad\quad\quad \left. + A^{\mathsf{H}} \left(z_{\mathrm{B}}^k - w_{\mathrm{B}}^k \right) \right)$
6: $\quad y_\ell^{k+1} = x^{k+1} + w_\ell^k$ $(\ell = 1, \ldots, S)$
7: $\quad y_{\mathrm{GS}}^{k+1} = x^{k+1} + w_{\mathrm{GS}}^k$
8: $\quad y_{\mathrm{B}}^{k+1} = A x^{k+1} + w_{\mathrm{B}}^k$
9: $\quad z_\ell^{k+1} = c_\ell \mathbf{1} + \mathrm{prox}_{\frac{q_\ell}{2\rho} g_\ell} \left(y_\ell^{k+1} - c_\ell \mathbf{1} \right)$
10: $\quad\quad (\ell = 1, \ldots, S)$
11: $\quad z_{\mathrm{GS},n}^{k+1} = \left[\mathrm{prox}_{\frac{\alpha}{2\rho} g_{\mathrm{GS}}} \left(y_{\mathrm{GS}}^{k+1} \right) \right]_n$ $(n = 1, \ldots, N)$
12: $\quad z_{\mathrm{B}}^{k+1} = \mathrm{prox}_{\frac{1}{2\rho} \chi_{\mathcal{B}}} \left(y_{\mathrm{B}}^{k+1} \right)$
13: $\quad w_\ell^{k+1} = y_\ell^{k+1} - z_\ell^{k+1}$ $(\ell = 1, \ldots, L)$
14: $\quad w_{\mathrm{GS}}^{k+1} = y_{\mathrm{GS}}^{k+1} - z_{\mathrm{GS}}^{k+1}$
15: $\quad w_{\mathrm{B}}^{k+1} = y_{\mathrm{B}}^{k+1} - z_{\mathrm{B}}^{k+1}$
16: **end for**
17: $\hat{s} = x^{K_{\mathrm{itr}}}$

D) IW-DGS

By using the weighting technique in [5, 13, 32], we here extend the DGS optimization into the W-DGS optimization so that we can use prior information. Moreover, we propose an iterative approach named IW-DGS, where we iteratively solve the W-DGS optimization with the updated parameters.

Assuming that the sparse regularizer $g_\ell(\cdot)$ is separable as $h_1(\cdot)$ or $h_2(\cdot)$, we extend the DGS optimization (12) to the W-DGS optimization given by

$$\underset{\boldsymbol{x}\in\mathbb{C}^{QN}}{\text{minimize}} \left\{ \sum_{\ell=1}^{S}\sum_{i=1}^{QN} q_{i,\ell}g_\ell\left(x_i - c_\ell\right) + \sum_{n=1}^{N}\alpha_n\|\boldsymbol{x}_n\|_2 \right\}$$

$$\text{subject to } \|\boldsymbol{r} - \boldsymbol{Ax}\|_2 \leq \varepsilon, \quad (38)$$

where $q_{i,\ell}, \alpha_n \, (> 0)$ are the parameters. When $q_{i,\ell} = q_\ell$ $(i = 1,\ldots,QN)$ and $\alpha_n = \alpha$ $(n = 1,\ldots,N)$, the W-DGS optimization (38) is equivalent to the DGS optimization (12). In the W-DGS optimization (38), we can use the different coefficients $q_{i,\ell}$ and α_n for each transmitted symbol s_i and transmitted subvector $\boldsymbol{s}_n$, respectively. We can thus tune these parameters if we have some prior information about $\boldsymbol{s}$, e.g. a tentative estimate of $\boldsymbol{s}$.

We provide the ADMM-based algorithm for the W-DGS optimization (38). Letting $\tilde{g}_{\text{GS},n}(\boldsymbol{z}_{\text{GS}}) = \|\boldsymbol{z}_{\text{GS},n}\|_2$ and $\tilde{f}(\boldsymbol{z}) = \sum_{\ell=1}^{S}\sum_{i=1}^{QN}q_{i,\ell}g_\ell(z_{i,\ell} - c_\ell) + \sum_{n=1}^{N}\alpha_n\tilde{g}_{\text{GS},n}(\boldsymbol{z}_{\text{GS}}) + \chi_{\mathcal{B}}(\boldsymbol{z}_{\text{B}})$, we can rewrite the W-DGS optimization problem (38) as

$$\underset{\substack{\boldsymbol{x}\in\mathbb{C}^{QN} \\ \boldsymbol{z}\in\mathbb{C}^{(S+1)QN+QM}}}{\text{minimize}} \quad \tilde{f}(\boldsymbol{z})$$

$$\text{subject to } \quad \boldsymbol{\Phi x} = \boldsymbol{z}, \quad (39)$$

where $\boldsymbol{z} = \begin{bmatrix} \boldsymbol{z}_1^\mathsf{T} \cdots \boldsymbol{z}_S^\mathsf{T} \boldsymbol{z}_{\text{GS}}^\mathsf{T} \boldsymbol{z}_{\text{B}}^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{C}^{(S+1)QN+QM}$, $\boldsymbol{z}_\ell = \begin{bmatrix} z_{1,\ell} \cdots z_{QN,\ell} \end{bmatrix}^\mathsf{T} \in \mathbb{C}^{QN}$ $(\ell = 1,\ldots,S)$, and $\boldsymbol{z}_{\text{GS}} = \begin{bmatrix} \boldsymbol{z}_{\text{GS},1}^\mathsf{T} \cdots \boldsymbol{z}_{\text{GS},N}^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{C}^{QN}$. The ADMM-based algorithm for (39) can be obtained by replacing the update of $\boldsymbol{z}_\ell^k$ and $\boldsymbol{z}_{\text{GS},n}^k$ in Algorithm 1 with

$$\boldsymbol{z}_\ell^{k+1} = \begin{bmatrix} c_\ell + \text{prox}_{\frac{q_{1,\ell}}{2\rho}g_\ell}\left(y_{1,\ell}^{k+1} - c_\ell\right) \\ \vdots \\ c_\ell + \text{prox}_{\frac{q_{CN,\ell}}{2\rho}g_\ell}\left(y_{CN,\ell}^{k+1} - c_\ell\right) \end{bmatrix}, \quad (40)$$

$$\boldsymbol{z}_{\text{GS},n}^{k+1} = \text{prox}_{\frac{\alpha_n}{2\rho}\tilde{g}_{\text{GS},n}}\left(\boldsymbol{y}_{\text{GS},n}^{k+1}\right) \quad (41)$$

$$= \begin{cases} \left(\left\|\boldsymbol{y}_{\text{GS},n}^{k+1}\right\|_2 - \frac{\alpha_n}{2\rho}\right)\frac{\boldsymbol{y}_{\text{GS},n}^{k+1}}{\left\|\boldsymbol{y}_{\text{GS},n}^{k+1}\right\|_2} & \left(\left\|\boldsymbol{y}_{\text{GS},n}^{k+1}\right\|_2 \geq \frac{\alpha_n}{2\rho}\right) \\ \boldsymbol{0}_Q & \left(\left\|\boldsymbol{y}_{\text{GS},n}^{k+1}\right\|_2 < \frac{\alpha_n}{2\rho}\right) \end{cases}, \quad (42)$$

respectively, where the vector $\boldsymbol{y}_{\text{GS},n}^{k+1} \in \mathbb{C}^Q$ is the n-th subvector of $\boldsymbol{y}_{\text{GS}}^{k+1} = [\boldsymbol{y}_{\text{GS},1}^{k+1}{}^\mathsf{T} \cdots \boldsymbol{y}_{\text{GS},N}^{k+1}{}^\mathsf{T}]^\mathsf{T}$.

We propose an iterative algorithm named IW-DGS in Algorithm 2, where we iteratively compute the solution of

---

Algorithm 2 IW-DGS

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

1: Initialize $q_{i,\ell} = q_\ell$ $(i = 1,\ldots,QN$ and $\ell = 1,\ldots,S)$ and $\alpha_n = \alpha$ $(n = 1,\ldots,N)$
2: **for** $t = 1$ to T **do**
3:     Fix $\rho, \alpha > 0$, $\boldsymbol{z}_1^\text{o},\ldots,\boldsymbol{z}_L^\text{o}, \boldsymbol{z}_{\text{GS}}^\text{o}, \boldsymbol{w}_1^\text{o},\ldots,\boldsymbol{w}_L^\text{o}, \boldsymbol{w}_{\text{GS}}^\text{o} \in \mathbb{C}^{QN}$, and $\boldsymbol{z}_\text{B}^\text{o}, \boldsymbol{w}_\text{B}^\text{o} \in \mathbb{C}^{QM}$
4:     **for** $k = 0$ to $K_{\text{itr}} - 1$ **do**
5:         $\boldsymbol{x}^{k+1} = \left((S+1)\boldsymbol{I}_{CN} + \boldsymbol{A}^\mathsf{H}\boldsymbol{A}\right)^{-1}$
6:             $\times\left(\sum_{\ell=1}^{S}\left(\boldsymbol{z}_\ell^k - \boldsymbol{w}_\ell^k\right) + \left(\boldsymbol{z}_{\text{GS}}^k - \boldsymbol{w}_{\text{GS}}^k\right)\right.$
7:                $\left.+\boldsymbol{A}^\mathsf{H}\left(\boldsymbol{z}_\text{B}^k - \boldsymbol{w}_\text{B}^k\right)\right)$
8:         $\boldsymbol{y}_\ell^{k+1} = \boldsymbol{x}^{k+1} + \boldsymbol{w}_\ell^k \, (\ell = 1,\ldots,S)$
9:         $\boldsymbol{y}_{\text{GS}}^{k+1} = \boldsymbol{x}^{k+1} + \boldsymbol{w}_{\text{GS}}^k$
10:         $\boldsymbol{y}_\text{B}^{k+1} = \boldsymbol{Ax}^{k+1} + \boldsymbol{w}_\text{B}^k$
11:         $\boldsymbol{z}_{i,\ell}^{k+1} = c_\ell + \text{prox}_{\frac{q_{i,\ell}}{2\rho}g_\ell}\left(\boldsymbol{y}_{i,\ell}^{k+1} - c_\ell\right)$
12:             $(i = 1,\ldots,QN$ and $\ell = 1,\ldots,S)$
13:         $\boldsymbol{z}_{\text{GS},n}^{k+1} = \text{prox}_{\frac{\alpha_n}{2\rho}\tilde{g}_{\text{GS},n}}\left(\boldsymbol{y}_{\text{GS},n}^{k+1}\right) \, (n = 1,\ldots,N)$
14:         $\boldsymbol{z}_\text{B}^{k+1} = \text{prox}_{\frac{1}{2\rho}\chi_{\mathcal{B}}}\left(\boldsymbol{y}_\text{B}^{k+1}\right)$
15:         $\boldsymbol{w}_\ell^{k+1} = \boldsymbol{y}_\ell^{k+1} - \boldsymbol{z}_\ell^{k+1} \, (\ell = 1,\ldots,L)$
16:         $\boldsymbol{w}_{\text{GS}}^{k+1} = \boldsymbol{y}_{\text{GS}}^{k+1} - \boldsymbol{z}_{\text{GS}}^{k+1}$
17:         $\boldsymbol{w}_\text{B}^{k+1} = \boldsymbol{y}_\text{B}^{k+1} - \boldsymbol{z}_\text{B}^{k+1}$
18:     **end for**
19:     $q_{i,\ell} = \dfrac{\left|x_i^{K_{\text{itr}}} - c_\ell\right|^{-1}}{\sum_{\ell'=1}^{S}\left|x_i^{K_{\text{itr}}} - c_{\ell'}\right|^{-1}}$
20:         $(i = 1,\ldots,QN$ and $\ell = 1,\ldots,S)$
21:     $\alpha_n = \dfrac{\left\|\boldsymbol{x}_n^{K_{\text{itr}}}\right\|_2^{-1}}{\sum_{n'=1}^{N}\left\|\boldsymbol{x}_{n'}^{K_{\text{itr}}}\right\|_2^{-1}}N\alpha \, (n = 1,\ldots,N)$
22: **end for**
23: $\hat{\boldsymbol{s}} = \boldsymbol{x}^{K_{\text{itr}}}$

---

the W-DGS optimization with the update of the parameters $q_{i,\ell}$ and α_n. In Fig. 2, we show the illustration of the proposed IW-DGS. At each outer iteration t, we update the parameters by using the tentative estimate at the previous iteration $t - 1$. In this paper, we consider the update given by

$$q_{i,\ell} = \frac{\left|\hat{s}_i^{\text{pre}} - c_\ell\right|^{-1}}{\sum_{\ell'=1}^{S}\left|\hat{s}_i^{\text{pre}} - c_{\ell'}\right|^{-1}}, \quad (43)$$

$$\alpha_n = \frac{\left\|\hat{\boldsymbol{s}}_n^{\text{pre}}\right\|_2^{-1}}{\sum_{n'=1}^{N}\left\|\hat{\boldsymbol{s}}_{n'}^{\text{pre}}\right\|_2^{-1}}N\alpha, \quad (44)$$

where we define $\hat{\boldsymbol{s}}^{\text{pre}} = \left[(\hat{\boldsymbol{s}}_1^{\text{pre}})^\mathsf{T} \cdots (\hat{\boldsymbol{s}}_N^{\text{pre}})^\mathsf{T}\right]^\mathsf{T} = [\hat{s}_1^{\text{pre}} \cdots \hat{s}_{QN}^{\text{pre}}]^\mathsf{T} \in \mathbb{C}^{QN}$ as the estimate of $\boldsymbol{s}$ at the previous iteration. The denominators in (43) and (44) play the role of normalization to satisfy $\sum_{\ell=1}^{S}q_{i,\ell} = 1$ for any $i = 1,\ldots,QN$ and $\sum_{n=1}^{N}\alpha_n = N\alpha$, respectively. By using (43) and (44), we can utilize the tentative estimate $\hat{\boldsymbol{s}}^{\text{pre}}$ as the prior information.
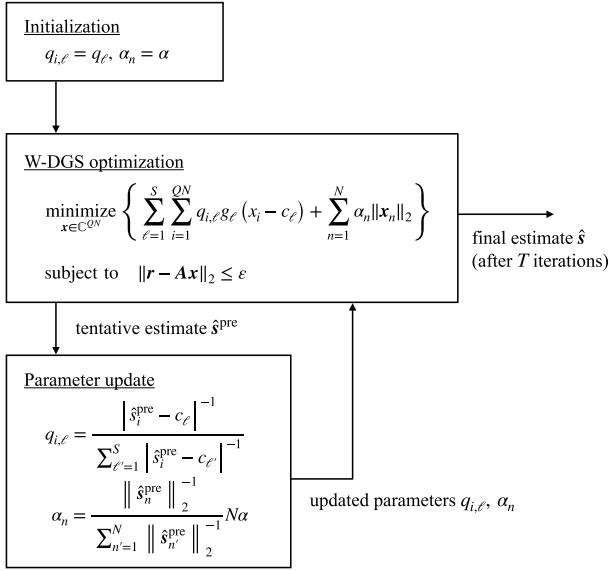
Fig. 2. Illustration of IW-DGS.

For example, the parameter $q_{i,\ell}$ becomes large when the corresponding tentative estimate $\hat{s}_i^{\mathrm{pre}}$ is close to c_ℓ. In this case, the new estimate of s_i also tends to be close to c_ℓ. Similarly, the parameter α_n becomes large when the norm of corresponding tentative estimate $\hat{s}_n^{\mathrm{pre}}$ is small. In this case, the new estimate of s_n also tends to become almost $\mathbf{o}_Q$.

E) Computational complexity

We here discuss the computational complexity of IW-DGS in Algorithm 2. The order of the complexity is dominated by the computation of the inverse matrix $((S+1)\boldsymbol{I}_{QN} + \boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})^{-1}$, which requires $\mathcal{O}(Q^3N^3)$ complexity in a direct calculation [33, Ch. 11]. However, by using the structure of $\boldsymbol{A}$, the order of the complexity can be reduced.

We firstly consider the computational complexity for $((S+1)\boldsymbol{I}_{QN} + \boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})^{-1}$ in the case of MU-MIMO OFDM systems (11) without precoding, i.e. $\boldsymbol{P} = \boldsymbol{I}_Q$. In this case, $\boldsymbol{\Theta} := (S+1)\boldsymbol{I}_{QN} + \boldsymbol{A}^{\mathsf{H}}\boldsymbol{A} \in \mathbb{C}^{QN \times QN}$ is composed of N^2 diagonal matrices as

$$
\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{\Theta}^{(1,1)} & \cdots & \boldsymbol{\Theta}^{(1,N)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Theta}^{(N,1)} & \cdots & \boldsymbol{\Theta}^{(N,N)} \end{bmatrix}, \tag{45}
$$

where $\boldsymbol{\Theta}^{(n_1,n_2)} = \mathrm{diag}(\theta_1^{(n_1,n_2)}, \ldots, \theta_Q^{(n_1,n_2)}) = \delta_{n_1 n_2}(S+1)$ $\boldsymbol{I}_Q + \sum_{m=1}^{M}(\boldsymbol{\Lambda}^{(m,n_1)})^{\mathsf{H}} \boldsymbol{\Lambda}^{(m,n_2)} \in \mathbb{C}^{Q \times Q}$ and $\delta_{n_1 n_2}$ denotes the Dirac delta. Since $\boldsymbol{\Theta}^{(n_1,n_2)}$ is a diagonal matrix as well as $\boldsymbol{\Lambda}^{(m,n_1)}$ and $\boldsymbol{\Lambda}^{(m,n_2)}$, each $\boldsymbol{\Theta}^{(n_1,n_2)}$ can be computed with $\mathcal{O}(QM)$. We define

$$
\tilde{\boldsymbol{\Theta}}_c = \begin{bmatrix} \theta_c^{(1,1)} & \cdots & \theta_c^{(1,N)} \\ \vdots & \ddots & \vdots \\ \theta_c^{(N,1)} & \cdots & \theta_c^{(N,N)} \end{bmatrix} \in \mathbb{C}^{N \times N} \tag{46}
$$

and its inverse matrix

$$
\tilde{\boldsymbol{\Omega}}_c = \begin{bmatrix} \omega_c^{(1,1)} & \cdots & \omega_c^{(1,N)} \\ \vdots & \ddots & \vdots \\ \omega_c^{(N,1)} & \cdots & \omega_c^{(N,N)} \end{bmatrix} \tag{47}
$$

$$
:= \tilde{\boldsymbol{\Theta}}_c^{-1} \in \mathbb{C}^{N \times N}, \tag{48}
$$

which correspond to the c-th subcarrier. The inverse matrix of $\boldsymbol{\Theta}$ in (45) can be written as

$$
\boldsymbol{\Theta}^{-1} = \begin{bmatrix} \boldsymbol{\Omega}^{(1,1)} & \cdots & \boldsymbol{\Omega}^{(1,N)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Omega}^{(N,1)} & \cdots & \boldsymbol{\Omega}^{(N,N)} \end{bmatrix}, \tag{49}
$$

where $\boldsymbol{\Omega}^{(n_1,n_2)} = \mathrm{diag}(\omega_1^{(n_1,n_2)}, \ldots, \omega_Q^{(n_1,n_2)}) \in \mathbb{C}^{Q \times Q}$. Since each $\tilde{\boldsymbol{\Omega}}_c \in \mathbb{C}^{N \times N}$ $(c = 1, \ldots, Q)$ in (48) can be obtained with $\mathcal{O}(N^3)$, the overall complexity for the inverse matrix $((S+1)\boldsymbol{I}_{QN} + \boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})^{-1}$ becomes $\mathcal{O}(QN^3)$, which is much lower than $\mathcal{O}(Q^3N^3)$ in the direct calculation. It should be noted that we can compute each $\tilde{\boldsymbol{\Omega}}_c$ in parallel. Moreover, the computation of the inverse matrix is required only once in the algorithm. Since the matrix-vector multiplication related to $\boldsymbol{A}$ requires $\mathcal{O}(Q^2MN)$, the overall complexity of IW-DGS with respect to the system size is $\mathcal{O}(Q^2MN + QN^3)$ in this case. Note that, in practice, we can compute the matrix-vector multiplication more efficiently by using the sparsity of $\boldsymbol{A}$, i.e. the number of non-zero elements in $\boldsymbol{A}$ is only QMN, whereas the number of all elements is Q^2MN.

Even when we use some precoding, we can also reduce the order of the complexity as long as the precoding matrix $\boldsymbol{P}$ is orthogonal, i.e. $\boldsymbol{P}\boldsymbol{P}^{\mathsf{H}} = \boldsymbol{I}_Q$. From the Sherman–Morrison–Woodbury formula [34], we have

$$
((S+1)\boldsymbol{I}_{QN} + \boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})^{-1} = \frac{1}{S+1}\boldsymbol{I}_{QN} - \frac{1}{(S+1)^2}\boldsymbol{A}^{\mathsf{H}}
$$
$$
\times \left(\boldsymbol{I}_{QM} + \frac{1}{S+1}\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}}\right)^{-1}\boldsymbol{A}. \tag{50}
$$

By using the orthogonality of $\boldsymbol{P}$, the matrix $\boldsymbol{\Xi} := \boldsymbol{I}_{QM} + \frac{1}{S+1}\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}} \in \mathbb{C}^{QM \times QM}$ in (50) can be written with M^2 diagonal matrices as

$$
\boldsymbol{\Xi} = \begin{bmatrix} \boldsymbol{\Xi}^{(1,1)} & \cdots & \boldsymbol{\Xi}^{(1,M)} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Xi}^{(M,1)} & \cdots & \boldsymbol{\Xi}^{(M,M)} \end{bmatrix}, \tag{51}
$$

where $\boldsymbol{\Xi}^{(m_1,m_2)}$ is given by $\boldsymbol{\Xi}^{(m_1,m_2)} = \delta_{m_1 m_2}\boldsymbol{I}_Q + (1/(S+1))$ $\sum_{m=1}^{M}\boldsymbol{\Lambda}^{(m,n_2)}\boldsymbol{P}(\boldsymbol{\Lambda}^{(m,n_1)}\boldsymbol{P})^{\mathsf{H}} = \delta_{m_1 m_2}\boldsymbol{I}_Q + (1/(S+1))\sum_{m=1}^{M}$ $\boldsymbol{\Lambda}^{(m,n_2)}(\boldsymbol{\Lambda}^{(m,n_1)})^{\mathsf{H}} \in \mathbb{C}^{Q \times Q}$ $(m_1, m_2 = 1, \ldots, M)$. Since each $\boldsymbol{\Xi}^{(m_1,m_2)}$ is a diagonal matrix, we can calculate the inverse matrix $\boldsymbol{\Xi}^{-1}$ with $\mathcal{O}(QM^3)$ in the same manner for $\boldsymbol{\Theta}$ in (45). Once we have $\boldsymbol{\Xi}^{-1} = (\boldsymbol{I}_{QM} + (1/(S+1))\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}})^{-1}$, we can update $\boldsymbol{x}^k$ only with some matrix-vector multiplications by using (50). Given that the matrix-vector multiplication requires $\mathcal{O}(Q^2MN)$, the overall complexity is given

Table 2. Convex optimization-based methods for discrete-valued vector reconstruction

Method	SOAV		SSR	SCSR			DGS (proposed)
Reference	[23, 24]	[10]	[30]	[13]	[12]	[30]	
Domain	real	real	real	complex	complex	complex	complex
Convexity	convex	convex	non-convex	convex	convex	non-convex	convex
Measurement model	i.i.d. etc.	MU-MIMO	i.i.d.	i.i.d. etc.	MU-MIMO	i.i.d.	MU-MIMO
Discreteness	✓	✓	✓	✓	✓	✓	✓
Group sparsity							✓

by $\mathcal{O}(Q^2MN + QM^3)$ in this case. One drawback of such approaches is that we require more number of matrix-vector multiplications compared to the case where we obtain the inverse matrix $((S+1)\boldsymbol{I}_{QN} + \boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})^{-1}$ in advance. However, when the complexity for $((S+1)\boldsymbol{I}_{QN} + \boldsymbol{A}^{\mathsf{H}}\boldsymbol{A})^{-1}$ is prohibitive, the approach with the Sherman–Morrison–Woodbury formula (50) would be a promising candidate.

F) Related work

We here discuss some related optimization problems proposed for the discrete-valued vector reconstruction. Table 2 summarizes the conventional optimization problems and the proposed DGS optimization.

The use of the discreteness in the optimization problem has been proposed as the SOAV optimization [23, 24], where the sum of ℓ_1 norms is used as the regularizer for the discrete-valued vector. The idea of the SOAV optimization has been extended to the discrete-valued vector reconstruction in the complex-valued domain [13]. Moreover, the extension to non-convex optimization has been proposed for both real- and complex-valued cases [30]. The non-convex optimization in the real-valued domain is called sum of sparse regularizers (SSR) optimization. Such non-convex optimization-based approaches can achieve better performance than the convex ones in some cases. However, the convergence of the algorithm is not guaranteed for the non-convex optimization, and the performance is sensitive to the choice of the parameters.

The works in [13, 23, 24, 30] do not focus on the signal detection in MU-MIMO systems, and hence the structure of the measurement matrix is different from the channel matrix considered in this paper. In [30], for example, the ideal i.i.d. Gaussian matrix is used. To evaluate the performance in the MU-MIMO signal detection, the convex SOAV optimization and SCSR optimization have been applied to the scenario in [10, 12], respectively. The results show that the use of the discreteness by the convex optimization problems is effective also in the MU-MIMO signal detection.

The main advantage of the proposed DGS optimization against the above methods is to utilize the group sparsity of the unknown vector, which is preferable for the signal detection in IoT environments considered in this paper. Moreover, unlike [30], the convergence of the proposed algorithm based on ADMM is guaranteed because the DGS optimization is convex.

IV. SIMULATION RESULTS

We evaluate the performance of the proposed signal detection methods via computer simulations. We assume the QPSK modulation with $1+j, -1+j, -1-j, 1-j$ and set $(c_1, c_2, c_3, c_4, c_5) = (0, 1+j, -1+j, -1-j, 1-j)$. The number of subcarriers is $Q = 64$ and the length of channel impulse response is $L = 10$. We use $g_1(\cdot) = h_1(\cdot)$ and $g_\ell(\cdot) = h_2(\cdot)$ ($\ell = 2, \ldots, 5$) as the sparse regularizers. In the simulations, the received signal $\boldsymbol{r}$ and the channel matrix $\boldsymbol{A}$ are scaled such that the scaled channel matrix $\tilde{\boldsymbol{A}}$ satisfies $\|\tilde{\boldsymbol{A}}\|_2 = 1$. We set the parameter ε in the optimization problems (12) and (38) as $\sqrt{QM\sigma_{\mathrm{v}}^2}$, where σ_{v}^2 denotes the noise variance after the scaling. The parameter of ADMM is set to $\rho = 1$. The number of inner iterations in ADMM is $K_{\mathrm{itr}} = 200$.

We investigate the initial parameters q_ℓ ($\ell = 1, \ldots, S$) and α of the proposed IW-DGS in Algorithm 2. From the symmetry of the QPSK constellation, we can set $q_2 = \cdots = q_5$. Hence, we firstly tune the parameter q_1 by evaluating the performance of the DGS optimization with $q_2 = \cdots = q_5 = (1 - q_1)/4$. Figure 3 shows the SER performance of the proposed IW-DGS with $\alpha = 20$ versus q_1. In the figure, T denotes the number of outer iterations in IW-DGS. We assume the MU-MIMO OFDM system without precoding ($\boldsymbol{P} = \boldsymbol{I}_Q$), where $(N, M) = (100, 25)$, $N_{\mathrm{act}} = 15$ and $E_{\mathrm{b}}/N_0 = 15$ dB. The figure shows that the best parameter in this case is $q_1 = 0$, which means that we do not have to promote the sparsity of the transmitted symbol vector by the first regularizer $\sum_{\ell=1}^{S} \sum_{i=1}^{QN} q_{i,\ell} g_\ell(x_i - c_\ell)$ in (38). This could be because the sparsity is sufficiently promoted by the second regularizer $\sum_{n=1}^{N} \alpha_n \|\boldsymbol{x}_n\|_2$ in (38) when $\alpha_n > 0$. We thus use $q_1 = 0$ and $q_2 = \cdots = q_5 = 0.25$ hereafter. Next, we examine the parameter α to control the balance between the two regularizers. Figure 4 shows the SER performance of IW-DGS with $q_1 = 0$ and $q_2 = \cdots = q_5 = 0.25$ versus α for $(N, M) = (100, 40)$, $N_{\mathrm{act}} = 25$ and $E_{\mathrm{b}}/N_0 = 15$ dB. We can see that the SER performance of IW-DGS is significantly improved as T increases. Note that $\alpha = 0$ means that we do not utilize the group sparsity of the transmitted signal vector at all, and hence the DGS optimization with $\alpha = 0$ corresponds to the conventional SCSR optimization [13]. The figure shows that we can significantly improve the SER performance by promoting the group sparsity with positive α. This also means that the proposed DGS optimization can achieve much better performance than the conventional SCSR optimization. From the figure, we fix $\alpha = 20$ in the following simulations.
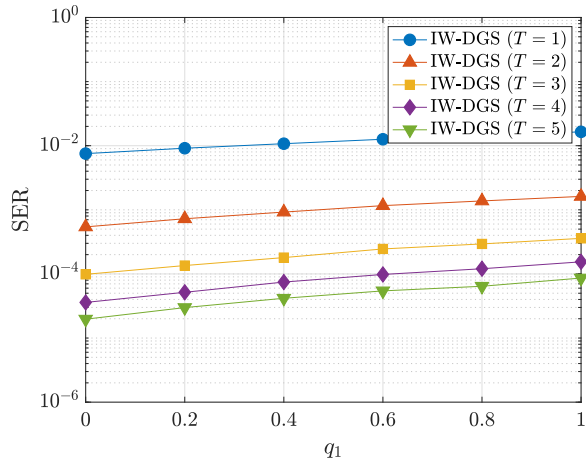
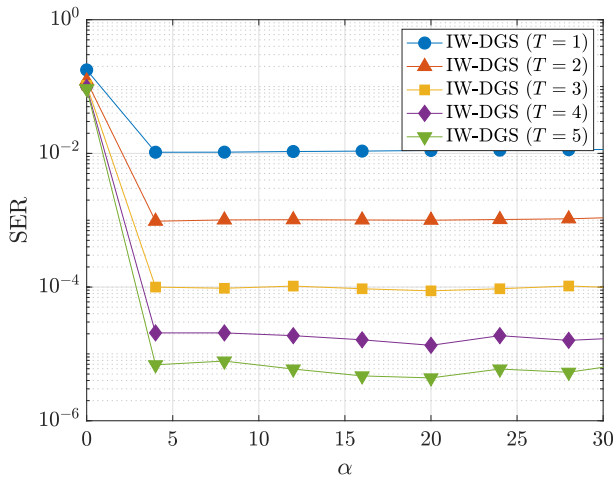**Fig. 3.** SER performance of IW-DGS with $\alpha = 20$ in MU-MIMO OFDM without precoding $((N, M) = (100, 25))$, $N_{\text{act}} = 15$, $E_{\text{b}}/N_{\text{o}} = 15$ dB).

Fig. 4. SER performance of IW-DGS in MU-MIMO OFDM without precoding $((N, M) = (100, 40))$, $N_{\text{act}} = 25$, $E_{\text{b}}/N_{\text{o}} = 15$ dB).

Figures 5 and 6 show the SER performance of IW-DGS versus $E_{\text{b}}/N_{\text{o}}$ in MU-MIMO OFDM systems without precoding. We have $(N, M) = (50, 25)$, $N_{\text{act}} = 10$ in Fig. 5, and $(N, M) = (100, 50)$, $N_{\text{act}} = 20$ in Fig. 6. In the figures, 'oracle ZF' denotes the oracle ZF method, which perfectly knows the activity of each IoT terminal. Specifically, the estimate of the oracle ZF method is given by $\hat{\boldsymbol{s}}_{\text{act}} = (\boldsymbol{A}_{\text{act}}^{\mathsf{H}} \boldsymbol{A}_{\text{act}})^{-1} \boldsymbol{A}_{\text{act}}^{\mathsf{H}} \boldsymbol{r}$, where $\hat{\boldsymbol{s}}_{\text{act}} \in \mathbb{C}^{QN_{\text{act}}}$ is the estimate of transmitted symbols of active IoT terminals and $\boldsymbol{A}_{\text{act}} \in \mathbb{C}^{QM \times QN_{\text{act}}}$ is the corresponding submatrix of the whole channel matrix $\boldsymbol{A}$. We can see that the performance of IW-DGS with $T = 10$ is close to that of the oracle ZF, especially for larger-scale problems. Moreover, the proposed IW-DGS can achieve better performance than the oracle ZF for high $E_{\text{b}}/N_{\text{o}}$.

We then evaluate the effect of the precoding matrix $\boldsymbol{P}$ in MU-MIMO OFDM systems. Figure 7 shows the SER performance of IW-DGS versus N. We have $M = 0.4N$, $N_{\text{act}} = 0.3N$, and $E_{\text{b}}/N_{\text{o}} = 15$ dB. In the figure, 'non-precoded' shows the SER performance when we do not use the precoding, i.e. $\boldsymbol{P} = \boldsymbol{I}_{Q}$. 'Hadamard' shows the performance with the precoding by the Hadamard matrix of order Q. 'DFT'
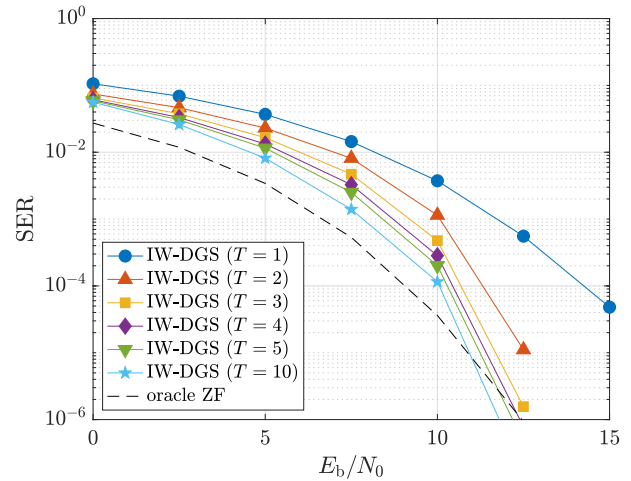

Fig. 5. SER performance of IW-DGS in MU-MIMO OFDM without precoding $((N, M) = (50, 25)$, $N_{\text{act}} = 10)$.
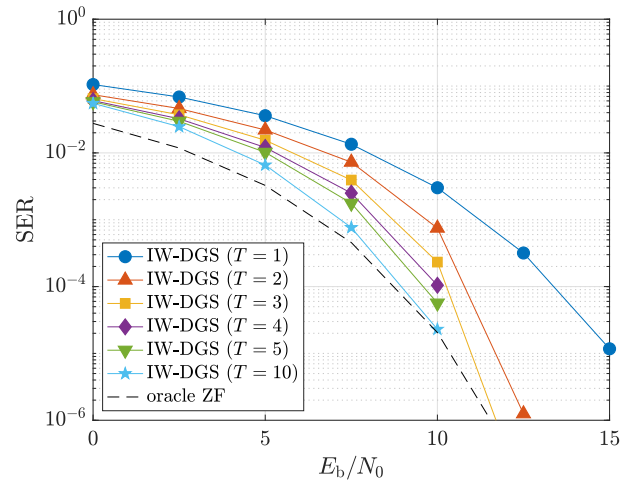

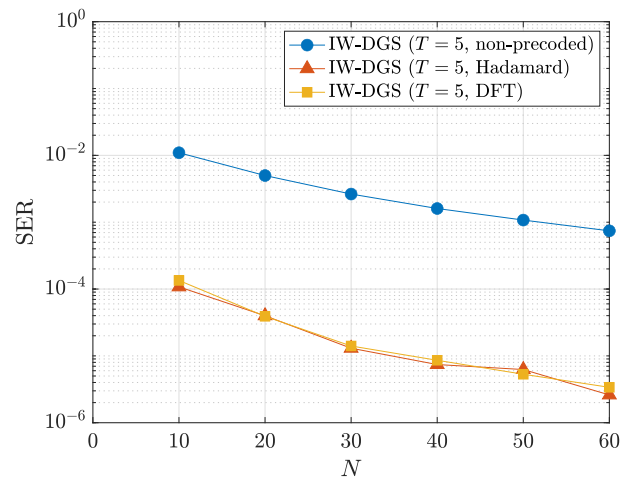**Fig. 6.** SER performance of IW-DGS in MU-MIMO OFDM without precoding $((N, M) = (100, 50)$, $N_{\text{act}} = 20)$.

Fig. 7. SER performance of IW-DGS versus N $(M = 0.4N$, $N_{\text{act}} = 0.3N$, $E_{\text{b}}/N_{\text{o}} = 15$ dB).
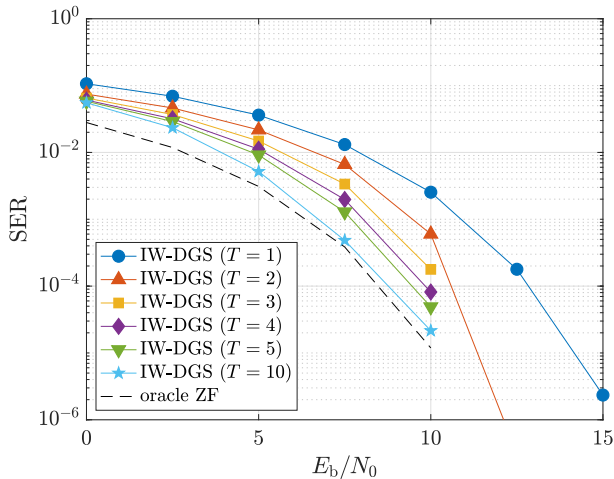
Fig. 8. SER performance of IW-DGS in MU-MIMO SC-CP ($(N, M) = (50, 25)$, $N_{\text{act}} = 10$).

shows the performance in the case with the DFT precoding, which is equivalent to MU-MIMO SC-CP systems. We can see that the performance in the precoded systems is much better than that in the non-precoded case. Even when $N = 10$, we can achieve the SER of 10^{-4} by using the precoding.

Finally, we show the SER performance versus $E_{\text{b}}/N_{\text{o}}$ in MU-MIMO SC-CP systems. Figure 8 shows the SER performance of IW-DGS in MU-MIMO SC-CP with $(N, M) = (50, 25)$ and $N_{\text{act}} = 10$. We can see that the proposed IW-DGS with $T = 10$ can also achieve almost the same performance as the oracle ZF in this case.

V. CONCLUSION

In this paper, we have proposed the signal detection method for the uplink overloaded MU-MIMO systems for IoT environments. First, we have shown that the MU-MIMO SC-CP system can be regarded as the MU-MIMO OFDM system with the DFT precoding. Since the SC-CP approach requires neither IDFT operation nor precoding matrix multiplication at the transmitter side, it is suited to the IoT data collection scenario, where cost, capability, and energy of the transmitter are extremely limited. Moreover, PAPR of SC-CP signals is lower than that of OFDM signals. We then have proposed the signal detection method named IW-DGS, where we iteratively solve the W-DGS optimization with updating the parameters in the objective function. The proposed IW-DGS can utilize both the discreteness and the group sparsity of the transmitted signal vector. By using the structure of the channel matrix, we can reduce the order of the computational complexity of the proposed IW-DGS. Simulation results show that IW-DGS can achieve good performance close to the oracle ZF method in the overloaded scenarios, where the number of receiving antennas is less than that of IoT terminals. Moreover, by using the precoding with the Hadamard matrix or the DFT matrix, we can significantly improve the SER performance of the

proposed method. Future work includes the extensive performance evaluation of the proposed method and extension to a non-convex optimization-based approach with the iterative weight update.

STATEMENT OF INTEREST

None.

REFERENCES

[1] Gupta, A.; Jha, R.K.: A survey of 5G network: architecture and emerging technologies. *IEEE Access*, **3** (2015), 1206–1232.

[2] Andrews, J.G.; Buzzi, S.; Choi, W.; Hanly, S.V.; Lozano, A.; Soong, A.C.K.; Zhang, J.C.: What will 5G be?. *IEEE J. Sel. Areas Commun.*, **32** (6) (2014), 1065–1082.

[3] Saxena, A.K.; Fijalkow, I.; Swindlehurst, A.L.: Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink. *IEEE Trans. Signal Process.*, **65** (17) (2017), 4624–4634.

[4] Wong, K.K.; Paulraj, A.; Murch, R.D.: Efficient high-performance decoding for overloaded MIMO antenna systems. *IEEE Trans. Wirel. Commun.*, **6** (5) (2007), 1833–1843.

[5] Hayakawa, R.; Hayashi, K.: Convex optimization-based signal detection for massive overloaded MIMO systems. *IEEE Trans. Wirel. Commun.*, **16** (11) (2017), 7080–7091.

[6] Takahashi, T.; Ibi, S.; Sampei, S.: Criterion of adaptively scaled belief for PDA in overloaded MIMO channels, in *Proc. 51st Asilomar Conference on Signals, Systems, and Computers*, October 2017, 1094–1098.

[7] Datta, T.; Srinidhi, N.; Chockalingam, A.; Rajan, B.S.: Low-complexity near-optimal signal detection in underdetermined large-MIMO systems, in *Proc. National Conference on Communications (NCC)*, February 2012, 1–5.

[8] Donoho, D.L.: Compressed sensing. *IEEE Trans. Inf. Theory*, **52** (4) (2006), 1289–1306.

[9] Hayashi, K.; Nagahara, M.; Tanaka, T.: A user's guide to compressed sensing for communications systems. *IEICE Trans. Commun.*, **E96-B** (3) (2013), 685–712.

[10] Hayashi, K.; Nakai, A.; Hayakawa, R.; Ha, S.: Uplink overloaded MU-MIMO OFDM signal detection methods using convex optimization, in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, November 2018, 1421–1427.

[11] Hayakawa, R.; Hayashi, K.: Discreteness-aware decoding for overloaded non-orthogonal STBCs via convex optimization. *IEEE Commun. Lett.*, **22** (10) (2018), 2080–2083.

[12] Hayashi, K.; Nakai-Kasai, A.; Hayakawa, R.: An overloaded SC-CP IoT signal detection method via sparse complex discrete-valued vector reconstruction, in *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, November 2019, 1473–1478.

[13] Hayakawa, R.; Hayashi, K.: Reconstruction of complex discrete-valued vector via convex optimization with sparse regularizers. *IEEE Access*, **6** (2018), 66499–66512.

[14] Liu, L.; Larsson, E.G.; Yu, W.; Popovski, P.; Stefanovic, C.; de Carvalho, E.: Sparse signal processing for grant-free massive connectivity: a future paradigm for random access protocols in the internet of things. *IEEE Signal Process. Mag.*, **35** (5) (2018), 88–99.

[15] Gabay, D.; Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Comput. Math. Appl.*, **2** (1) (1976), 17–40.

[16] Eckstein, J.; Bertsekas, D.P.: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.*, **55** (1) (1992), 293–318.

[17] Combettes, P.L.; Pesquet, J.-C.: Proximal Splitting Methods in Signal Processing, in Fixed-Point Algorithms for Inverse Problems in Science and Engineering ser. Springer Optimization and Its Applications, Springer, vol. **49**, New York, NY, 2011.185–212

[18] Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found Trends Mach. Learn.*, **3** (1) (2011), 1–122.

[19] Li, L.; Wang, X.; Wang, G.: Alternating direction method of multipliers for separable convex optimization of real functions in complex variables. *Math. Probl. Eng.*, **2015** (2015), 1–14.

[20] Wang, Z.; Giannakis, G.B.: Wireless multicarrier communications. *IEEE Signal Process. Mag.*, **17** (3) (2000), 29–48.

[21] Tan, P.H.; Rasmussen, L.K.; Lim, T.J.: Constrained maximum-likelihood detection in CDMA. *IEEE Trans. Commun.*, **49** (1) (2001), 142–153.

[22] Yener, A.; Yates, R.D.; Ulukus, S.: CDMA multiuser detection: a nonlinear programming approach. *IEEE Trans. Commun.*, **50** (6) (2002), 1016–1024.

[23] Aïssa-El-Bey, A.; Pastor, D.; Sbaï, S.M.A.; Fadlallah, Y.: Sparsity-based recovery of finite alphabet solutions to underdetermined linear systems. *IEEE Trans. Inf. Theory*, **61** (4) (2015), 2008–2018.

[24] Nagahara, M.: Discrete signal reconstruction by sum of absolute values. *IEEE Signal Process. Lett.*, **22** (10) (2015), 1575–1579.

[25] Yuan, M.; Lin, Y.: Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **68** (1) (2006), 49–67.

[26] Meier, L.; Geer, S.V.D.; Bühlmann, P.: The group lasso for logistic regression. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, **70** (1) (2008), 53–71.

[27] Stojnic, M.; Parvaresh, F.; Hassibi, B.: On the reconstruction of block-sparse signals with an optimal number of measurements. *IEEE Trans. Signal Process.*, **57** (8) (2009), 3075–3085.

[28] Messai, M.; Amis, K.; Guilloud, F.; Aïssa-El-Bey, A.: Reconstruction of finite-alphabet block-sparse signals from MAP support detection. *IEEE Access*, **7** (2019), 57996–58009.

[29] Messai, M.; Aïssa-El-Bey, A.; Amis, K.; Guilloud, F.: Iteratively reweighted two-stage LASSO for block-sparse signal recovery under finite-alphabet constraints. *Signal. Process.*, **157** (2019), 73–77.

[30] Hayakawa, R.; Hayashi, K.: Discrete-valued vector reconstruction by optimization with sum of sparse regularizers, in *Proc. 27th European Signal Processing Conference (EUSIPCO)*, September 2019, 1–5.

[31] Brandwood, D.H.: A complex gradient operator and its application in adaptive array theory. *IEE Proc. F - Commun. Radar Signal Process.*, **130** (1) (1983), 11–16.

[32] Candès, E.J.; Wakin, M.B.; Boyd, S.P.: Enhancing sparsity by reweighted ℓ_1 minimization. *J Fourier Anal. Appl.*, **14** (5) (2008), 877–905.

[33] Boyd, S.; Vandenberghe, L.: Introduction to Applied Linear Algebra: Vectors Matrices and Least Squares, Cambridge University Press, Cambridge, UK; New York, NY, 2018.

[34] Hager, W.: Updating the inverse of a matrix. *SIAM Rev.*, **31** (2) (1989), 221–239.

Ryo Hayakawa received the bachelor's degree in engineering, the master's degree in informatics, and the Ph.D. degree in informatics from Kyoto University, Kyoto, Japan, in 2015, 2017, and 2020, respectively. He is currently an Assistant Professor at the Graduate School of Engineering Science, Osaka University. He was a Research Fellow (DC1) of the Japan Society for the Promotion of Science (JSPS) from 2017 to 2020. He received the 33rd Telecom System Technology Student Award, APSIPA ASC 2019 Best Special Session Paper Nomination Award, and the 16th IEEE Kansai Section Student Paper Award. His research interests include signal processing and wireless communication. He is a member of IEEE and IEICE.

Ayano Nakai-Kasai received the B.E. and M.E. degrees from Kyoto University, Kyoto, Japan, in 2016 and 2018, respectively, where she is currently pursuing the Ph.D. degree with the Department of Systems Science, Graduate School of Informatics. Her research interests include signal processing and wireless communication. She received the Young Researchers' Award from the Institute of Electronics, Information and Communication Engineers in 2018 and APSIPA ASC 2019 Best Special Session Paper Nomination Award. She is a member of IEEE and IEICE.

Kazunori Hayashi is currently a Professor at the Center for Innovative Research and Education in Data Science/Graduate School of Informatics, Kyoto University. He received the B.E., M.E., and Ph.D. degrees in communication engineering from Osaka University, Osaka, Japan, in 1997, 1999, and 2002, respectively. He was an Assistant Professor from 2002 to 2007, an Associate Professor from 2007 to 2017 at the Graduate School of Informatics, Kyoto University, and a Professor from 2017 to 2020 at the Graduate School of Engineering, Osaka City University. His research interests include statistical signal processing for communication systems. He received the ICF Research Award from the KDDI Foundation in 2008, the IEEE Globecom 2009 Best Paper Award, the IEICE Communications Society Best Paper Award in 2010, the WPMC'11 Best Paper Award, the Telecommunications Advancement Foundation Award in 2012, the IEICE Communications Society Best Tutorial Paper Award in 2013, and APSIPA ASC Best Special Session Paper Nomination Award in 2019. He is a member of IEEE, IEICE, and ISCIE.