

Convex Optimization Based Signal Detection for Massive Overloaded MIMO Systems

Ryo Hayakawa, *Student Member, IEEE*, and Kazunori Hayashi, *Member, IEEE*

Abstract—This paper proposes signal detection schemes for massive multiple-input multiple-output (MIMO) systems, where the number of receive antennas is less than that of transmitted streams. Assuming practical baseband digital modulation, and taking advantage of the discreteness of transmitted symbols, we formulate the signal detection problem as a convex optimization problem, called sum-of-absolute-value (SOAV) optimization. Moreover, we extend the SOAV optimization into the weighted-SOAV (W-SOAV) optimization and propose an iterative approach to solve the W-SOAV optimization with updating the weights in the objective function. Furthermore, for coded MIMO systems, we also propose a joint detection and decoding scheme, where log likelihood ratios (LLRs) of transmitted symbols are iteratively updated between the MIMO detector and the channel decoder. In addition, a theoretical performance analysis is provided in terms of the upper bound of the size of the estimation error obtained with the W-SOAV optimization. Simulation results show that the bit error rate (BER) performance of the proposed scheme is better than that of conventional schemes, especially in large-scale overloaded MIMO systems.

Index Terms—massive MIMO, overloaded MIMO, sum-of-absolute-value optimization, proximal splitting methods, restricted isometry property.

I. INTRODUCTION

MULTIPLE-INPUT multiple-output (MIMO) communications use multiple antennas at both transmitter and receiver to achieve high spectral efficiency. As the required data rate and throughput have been significantly increasing, massive MIMO using tens or hundreds of antennas are gathering attention as one of key technologies in the 5th generation (5G) mobile communications systems [1], [2]. Since the computational complexity of MIMO signal detection generally increases along with the number of antennas, low-complexity signal detection schemes will be required for massive MIMO systems. Possible candidates might be linear signal detection schemes, such as the zero forcing (ZF) and the minimum mean-square-error (MMSE) detection methods, however, the performance of linear detectors is considerably poor compared to that of optimal maximum likelihood (ML) detection, and thus some non-linear detection schemes have been investigated to achieve better performance with reasonable computational complexity. The likelihood ascent search (LAS) [3], [4] and the reactive tabu search (RTS) [5], [6] are based on the local

neighborhood search of likelihood. The graph-based iterative Gaussian detector (GIGD) [7] is known as a low complexity scheme built upon belief propagation (BP). With the turbo principle as used in turbo-Bell Laboratories layered space time (BLAST) [8], the BP-based detection can be jointly performed with the decoding of channel codes such as low density parity check (LDPC) code [9], [10], [11]. It has been shown that this joint approach has better performance compared to the individual detection and decoding [12].

In some MIMO systems, sufficient number of receive antennas may not be available due to the limited size, weight, cost and/or power consumption of the receiver. Such MIMO systems, where the number of receive antennas is less than that of transmitted streams, are known as overloaded (or underdetermined) MIMO systems. As one of signal detection schemes for overloaded MIMO systems, the slab-sphere decoding (SSD) [13] based on ML detection has been proposed by extending the idea of sphere decoding [14]. On the other hand, pre-voting cancellation (PVC) [15] technique divides the fat channel matrix into a square matrix and the remaining matrix, and, for all possible candidates of the signals corresponding to the remaining matrix, signals corresponding to the square matrix are exhaustively estimated by using a conventional signal detection scheme for non-overloaded MIMO systems, and the one having best likelihood is selected as the final estimate of the transmitted signals. Moreover, to reduce the computational complexity, the integration of SSD and PVC has been considered in [16], [17]. For massive overloaded MIMO systems, however, these schemes are not feasible because their complexities are still too high. In addition, most of low-complexity detection schemes for non-overloaded massive MIMO systems, including LAS, RTS, and GIGD¹, have very poor performance in overloaded scenarios. While the signal detection scheme for massive overloaded MIMO has been hardly discussed in the literature, one of a few exceptions is the enhanced reactive tabu search (ERTS) proposed in [18]. ERTS is an extension of RTS and employs RTS iteratively while randomly varying the initial point of the search until a certain condition is satisfied. It is shown in [18] that ERTS can achieve comparable performance to the optimal ML detection with affordable computational complexity for overloaded MIMO systems with around 30 transmit antennas. If the number of antennas further increases, however, ERTS requires prohibitive computational complexity to achieve such performance because the required number of RTSs signifi-

¹GIGD works well for overloaded MIMO systems where the difference between the number of transmitted streams and receive antennas is small.

Manuscript received December 7, 2016; revised April 22, 2017.
Ryo Hayakawa is with the Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan (e-mail: rhayakawa@sys.i.kyoto-u.ac.jp).

Kazunori Hayashi is with the Graduate School of Engineering, Osaka City University, Osaka 558-8585, Japan (e-mail: kazunori@eng.osaka-cu.ac.jp).

This work has been partially presented in EUSIPCO 2016, Budapest, Hungary, August 29–September 2, 2016.

cantly increases. Other exceptions will be the ℓ_1 minimization-based method [19] and the quadratic programming-based method [20] for massive overloaded MIMO systems, and the quadratic programming-based method has been shown to have better performance than the ℓ_1 minimization-based method.

In this paper, based on the preliminary conference paper [21], we propose a low-complexity signal detection scheme for massive overloaded MIMO systems, which employs the fact that transmitted symbols in practical digital communications are discrete-valued. Using the discreteness and the idea of the sum-of-absolute-value (SOAV) optimization [22], [23], we formulate the signal detection problem as a convex optimization problem. The optimization problem can be efficiently solved with proximal splitting methods [24] even for underdetermined systems. Moreover, we extend the SOAV optimization to the weighted-SOAV (W-SOAV) optimization, where the prior information on transmitted symbols can be used, and propose an iterative approach, referred to as iterative weighted-SOAV (IW-SOAV). IW-SOAV calculates log likelihood ratios (LLRs) of transmitted symbols by using the estimate in the previous iteration, and utilizes them as the prior information. IW-SOAV can detect the transmitted signals with low computational complexity because the W-SOAV optimization problem can also be efficiently solved with proximal splitting methods. Furthermore, for coded MIMO systems, we propose a joint MIMO signal detection and channel decoding scheme, where LLRs of transmitted symbols are iteratively updated between the W-SOAV detector and the channel decoder. In addition, we analytically evaluate the error of the W-SOAV optimization by using the restricted isometry property (RIP) [25]. We derive an upper bound for the size of the error vector between the true transmitted vector and its estimate obtained with the W-SOAV optimization. Our analysis also provides a sufficient condition for the exact reconstruction with the W-SOAV optimization in the noise-free case. Simulation results show that IW-SOAV has better bit error rate (BER) performance than conventional signal detection schemes in large-scale overloaded MIMO systems. For LDPC coded MIMO systems, the joint detection and decoding can achieve much better performance than the individual detection and decoding.

Major differences against the conference version [21] include the joint processing of the overloaded massive MIMO signal detection and channel decoding, where improved weights in the W-SOAV optimization are newly introduced by using the LLRs of the transmitted symbols as the prior information. Furthermore, the theoretical performance analysis of the W-SOAV optimization based on the RIP is also newly added contribution in this paper.

The rest of the paper is organized as follows: In Section II, we introduce the system model of the overloaded MIMO system without channel coding and propose a signal detection scheme named IW-SOAV. We consider the coded case and explain the integrated scheme with IW-SOAV and the channel decoding in Section III. In Section IV, the theoretical performance analysis of the W-SOAV optimization by using RIP is provided. Section V gives some simulation results to demonstrate the performance of the proposed scheme. Finally,

Section VI presents some conclusions.

In the rest of the paper, we use the following notations: We denote the set of all real numbers by \mathbb{R} and the set of all complex numbers by \mathbb{C} . $\text{Re}\{\cdot\}$ and $\text{Im}\{\cdot\}$ indicate the real part and the imaginary part, respectively. Superscripts $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and the Hermitian transpose, respectively. We represent the imaginary unit by j , the identity matrix by \mathbf{I} , a vector whose elements are all 1 by $\mathbf{1}$, and a vector whose elements are all 0 by $\mathbf{0}$. For a vector $\mathbf{a} = [a_1 \cdots a_N]^T \in \mathbb{R}^N$, we define the ℓ_1 , ℓ_2 , and ℓ_∞ norms of \mathbf{a} as $\|\mathbf{a}\|_1 = \sum_{i=1}^N |a_i|$, $\|\mathbf{a}\|_2 = \sqrt{\sum_{i=1}^N a_i^2}$, and $\|\mathbf{a}\|_\infty = \max_{i \in \{1, \dots, N\}} |a_i|$, respectively. For an index set $\mathcal{I} \subset \{1, \dots, N\}$, $\mathbf{a}_{\mathcal{I}} \in \mathbb{R}^N$ is defined by

$$[\mathbf{a}_{\mathcal{I}}]_i = \begin{cases} a_i & (i \in \mathcal{I}) \\ 0 & (i \notin \mathcal{I}) \end{cases}, \quad (1)$$

where $[\cdot]_i$ denotes the i th element of the vector. $|\mathcal{I}|$ and $\mathcal{I}^c = \{1, \dots, N\} \setminus \mathcal{I}$ represent the cardinality of \mathcal{I} and the complement set of \mathcal{I} , respectively. We denote the Euclidean inner product by $\langle \cdot, \cdot \rangle$ and the sign function by $\text{sgn}(\cdot)$.

II. PROPOSED SIGNAL DETECTION SCHEME FOR OVERLOADED MIMO WITHOUT CHANNEL CODING

A. System model

We consider a MIMO system with n transmit antennas and m ($< n$) receive antennas. For simplicity, precoding is not considered and the number of transmitted streams is assumed to be equal to that of transmit antennas. The transmitted signal vector $\tilde{\mathbf{s}} = [\tilde{s}_1 \cdots \tilde{s}_n]^T \in \tilde{\mathbb{S}}^n$ is composed of symbols transmitted from the n transmit antennas, where \tilde{s}_j ($j = 1, \dots, n$) denotes the symbol sent from the j th transmit antenna and $\tilde{\mathbb{S}}$ is the set of symbol alphabets. The received signal vector $\tilde{\mathbf{y}} = [\tilde{y}_1 \cdots \tilde{y}_m]^T \in \mathbb{C}^m$, where \tilde{y}_i ($i = 1, \dots, m$) denotes the signal received at the i th receive antenna, is given by

$$\tilde{\mathbf{y}} = \tilde{\mathbf{H}}\tilde{\mathbf{s}} + \tilde{\mathbf{v}}, \quad (2)$$

where

$$\tilde{\mathbf{H}} = \begin{bmatrix} \tilde{h}_{1,1} & \cdots & \tilde{h}_{1,n} \\ \vdots & \ddots & \vdots \\ \tilde{h}_{m,1} & \cdots & \tilde{h}_{m,n} \end{bmatrix} \in \mathbb{C}^{m \times n} \quad (3)$$

is a flat fading channel matrix and $\tilde{h}_{i,j}$ represents the channel gain from the j th transmit antenna to the i th receive antenna. $\tilde{\mathbf{v}} \in \mathbb{C}^m$ is the circular complex Gaussian noise vector with zero mean and the covariance matrix of $\sigma_v^2 \mathbf{I}$. The complex signal model (2) can be rewritten by the real signal model as

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{v}, \quad (4)$$

where

$$\mathbf{y} = \begin{bmatrix} \text{Re}\{\tilde{\mathbf{y}}\} \\ \text{Im}\{\tilde{\mathbf{y}}\} \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \text{Re}\{\tilde{\mathbf{H}}\} & -\text{Im}\{\tilde{\mathbf{H}}\} \\ \text{Im}\{\tilde{\mathbf{H}}\} & \text{Re}\{\tilde{\mathbf{H}}\} \end{bmatrix}, \\ \mathbf{s} = \begin{bmatrix} \text{Re}\{\tilde{\mathbf{s}}\} \\ \text{Im}\{\tilde{\mathbf{s}}\} \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \text{Re}\{\tilde{\mathbf{v}}\} \\ \text{Im}\{\tilde{\mathbf{v}}\} \end{bmatrix}. \quad (5)$$

B. SOAV optimization

The SOAV optimization [22], [23] is a technique to reconstruct an unknown discrete-valued vector $\mathbf{x} = [x_1 \cdots x_N]^T \in \{c_1, \dots, c_B\}^N \subset \mathbb{R}^N$ from its possibly underdetermined linear measurements $\boldsymbol{\eta} = \mathbf{A}\mathbf{x}$, where $\mathbf{A} \in \mathbb{R}^{M \times N}$. If we assume $\Pr(x_j = c_b) = 1/B$ ($b = 1, \dots, B$) for all x_j ($j = 1, \dots, N$), each of $\mathbf{x} - c_1\mathbf{1}, \dots, \mathbf{x} - c_B\mathbf{1}$ has approximately N/B zero elements. Based on this property and the idea of the ℓ_1 optimization in compressed sensing [26], [27], the SOAV optimization solves

$$\underset{\mathbf{u} \in \mathbb{R}^N}{\text{minimize}} \quad \frac{1}{B} \sum_{b=1}^B \|\mathbf{u} - c_b\mathbf{1}\|_1 \quad \text{subject to} \quad \boldsymbol{\eta} = \mathbf{A}\mathbf{u} \quad (6)$$

in order to reconstruct \mathbf{x} from $\boldsymbol{\eta}$ and \mathbf{A} .

C. Signal detection via SOAV optimization

In digital MIMO communications systems, the transmitted signal vector \mathbf{s} is commonly discrete-valued and the received signal vector \mathbf{y} can be regarded as its linear observations if the noise can be ignored. For simplicity, if we employ quadrature phase shift keying (QPSK), then $\mathbb{S} = \{1 + j, -1 + j, -1 - j, 1 - j\}$, $\mathbb{E}[\tilde{\mathbf{s}}] = \mathbf{0}$, and $\mathbb{E}[\tilde{\mathbf{s}}\tilde{\mathbf{s}}^H] = 2\mathbf{I}$. Since each element of \mathbf{s} is 1 or -1 for this case, we can formulate the signal detection problem as the SOAV optimization, i.e.,

$$\underset{\mathbf{z} \in \mathbb{R}^{2n}}{\text{minimize}} \quad \left(\frac{1}{2} \|\mathbf{z} - \mathbf{1}\|_1 + \frac{1}{2} \|\mathbf{z} + \mathbf{1}\|_1 \right) \\ \text{subject to} \quad \mathbf{y} = \mathbf{H}\mathbf{z}. \quad (7)$$

Since the received signal vector \mathbf{y} contains the additive noise as in (4), we modify the optimization problem as

$$\hat{\mathbf{s}} = \underset{\mathbf{z} \in \mathbb{R}^{2n}}{\text{argmin}} \quad \left(\frac{1}{2} \|\mathbf{z} - \mathbf{1}\|_1 + \frac{1}{2} \|\mathbf{z} + \mathbf{1}\|_1 \right. \\ \left. + \frac{\alpha}{2} \|\mathbf{y} - \mathbf{H}\mathbf{z}\|_2^2 \right), \quad (8)$$

by using the idea of the ℓ_1 - ℓ_2 optimization [27]. Here, the constant $\alpha > 0$ is the parameter to control the balance between the terms $\frac{1}{2} \|\mathbf{z} - \mathbf{1}\|_1 + \frac{1}{2} \|\mathbf{z} + \mathbf{1}\|_1$ and $\frac{\alpha}{2} \|\mathbf{y} - \mathbf{H}\mathbf{z}\|_2^2$. If α is large, the effect of $\frac{\alpha}{2} \|\mathbf{y} - \mathbf{H}\mathbf{z}\|_2^2$ will be dominant compared to that of $\frac{1}{2} \|\mathbf{z} - \mathbf{1}\|_1 + \frac{1}{2} \|\mathbf{z} + \mathbf{1}\|_1$, and vice versa. In general, the optimal value of α is unknown in advance. For a fixed $\alpha > 0$, the solution of (8) can be obtained with the following theorem [24].

Theorem 1. Let $\phi_1, \phi_2 : \mathbb{R}^N \rightarrow (-\infty, \infty]$ be lower semicontinuous convex functions and $(\text{ri dom } \phi_1) \cap (\text{ri dom } \phi_2) \neq \emptyset$, where ri and dom denote the relative interior of the set and the domain of the function, respectively. In addition, $\phi_1(\mathbf{z}) + \phi_2(\mathbf{z}) \rightarrow \infty$ as $\|\mathbf{z}\|_2 \rightarrow \infty$ is assumed. A sequence $\{\mathbf{z}_k\}$ ($k = 0, 1, \dots$) converging to the solution of

$$\underset{\mathbf{z} \in \mathbb{R}^N}{\text{minimize}} \quad (\phi_1(\mathbf{z}) + \phi_2(\mathbf{z})) \quad (9)$$

can be obtained by using the following Douglas-Rachford algorithm.

Algorithm 1. (Douglas-Rachford algorithm)

1) Fix $\varepsilon \in (0, 1)$, $\gamma > 0$, and $\mathbf{r}_0 \in \mathbb{R}^N$.

2) For $k = 0, 1, 2, \dots$, iterate

$$\begin{cases} \mathbf{z}_k = \text{prox}_{\gamma\phi_1}(\mathbf{r}_k) \\ \theta_k \in [\varepsilon, 2 - \varepsilon] \\ \mathbf{r}_{k+1} = \mathbf{r}_k + \theta_k (\text{prox}_{\gamma\phi_2}(2\mathbf{z}_k - \mathbf{r}_k) - \mathbf{z}_k). \end{cases}$$

The Douglas-Rachford algorithm is one of the proximal splitting methods [24], which can solve the optimization problem with the form of (9) by using the proximity operator. For a lower semicontinuous convex function $\phi : \mathbb{R}^N \rightarrow \mathbb{R}$, the proximity operator is defined as

$$\text{prox}_{\phi}(\mathbf{z}) = \underset{\mathbf{u} \in \mathbb{R}^N}{\text{argmin}} \quad \left(\phi(\mathbf{u}) + \frac{1}{2} \|\mathbf{z} - \mathbf{u}\|_2^2 \right), \quad (10)$$

which is an extension of the projection onto nonempty closed convex sets for the convex function. In fact, for the indicator function $\iota_C(\mathbf{z})$ ($\iota_C(\mathbf{z}) = 0$ if $\mathbf{z} \in C$, and $\iota_C(\mathbf{z}) = \infty$ otherwise) with such a convex set C , $\text{prox}_{\iota_C}(\mathbf{z})$ is the projection of \mathbf{z} onto C . Although the choice of the parameters $\varepsilon, \gamma, \mathbf{r}_0$, and θ_k may affect the convergence speed of the algorithm, the limit of the sequence $\{\mathbf{z}_k\}$ is identical theoretically for any possible values of the parameters.

In order to apply the theorem to our problem, we rewrite (8) as

$$\underset{\mathbf{z} \in \mathbb{R}^{2n}}{\text{minimize}} \quad (f(\mathbf{z}) + g(\mathbf{z})), \quad (11)$$

where $f(\mathbf{z}) = \|\mathbf{z} - \mathbf{1}\|_1/2 + \|\mathbf{z} + \mathbf{1}\|_1/2$ and $g(\mathbf{z}) = \alpha \|\mathbf{y} - \mathbf{H}\mathbf{z}\|_2^2/2$. Note that $f(\mathbf{z})$ and $g(\mathbf{z})$ are lower semicontinuous convex functions due to the continuity and the convexity of ℓ_1 and ℓ_2 norms. Moreover, we have $(\text{ri dom } f) \cap (\text{ri dom } g) = (\text{ri } \mathbb{R}^{2n}) \cap (\text{ri } \mathbb{R}^{2n}) = \mathbb{R}^{2n} \neq \emptyset$ and $f(\mathbf{z}) + g(\mathbf{z}) \rightarrow \infty$ as $\|\mathbf{z}\|_2^2 \rightarrow \infty$. Thus, we can calculate the solution of (8) or (11) by using Algorithm 1 with $\phi_1(\mathbf{z}) = f(\mathbf{z})$ and $\phi_2(\mathbf{z}) = g(\mathbf{z})$. The proximity operators of $\gamma f(\mathbf{z})$ and $\gamma g(\mathbf{z})$ can be obtained as

$$[\text{prox}_{\gamma f}(\mathbf{z})]_j = \begin{cases} z_j + \gamma & (z_j < -1 - \gamma) \\ -1 & (-1 - \gamma \leq z_j < -1) \\ z_j & (-1 \leq z_j \leq 1) \\ 1 & (1 \leq z_j < 1 + \gamma) \\ z_j - \gamma & (1 + \gamma \leq z_j) \end{cases}, \quad (12)$$

and

$$\text{prox}_{\gamma g}(\mathbf{z}) = (\mathbf{I} + \alpha\gamma\mathbf{H}^T\mathbf{H})^{-1}(\mathbf{z} + \alpha\gamma\mathbf{H}^T\mathbf{y}), \quad (13)$$

respectively, where z_j indicates the j th element of \mathbf{z} . Note that $[\text{prox}_{\gamma f}(\mathbf{z})]_j$ is a function, which depends only on z_j , as shown in Fig. 1. By solving (8) via the Douglas-Rachford algorithm with $\text{prox}_{\gamma f}$ and $\text{prox}_{\gamma g}$, the estimate of the transmitted signal vector \mathbf{s} can be obtained.

The computational complexity of the algorithm is $O(n^3)$, which is dominated by the matrix inversion $(\mathbf{I} + \alpha\gamma\mathbf{H}^T\mathbf{H})^{-1}$ in (13). Note that the calculation of the inversion is required only once, and thus the corresponding computational cost does not grow with the number of iterations in the algorithm. If the ratio m/n is fixed, the computational complexity is the same order as that of linear MMSE detection for the overloaded scenario.

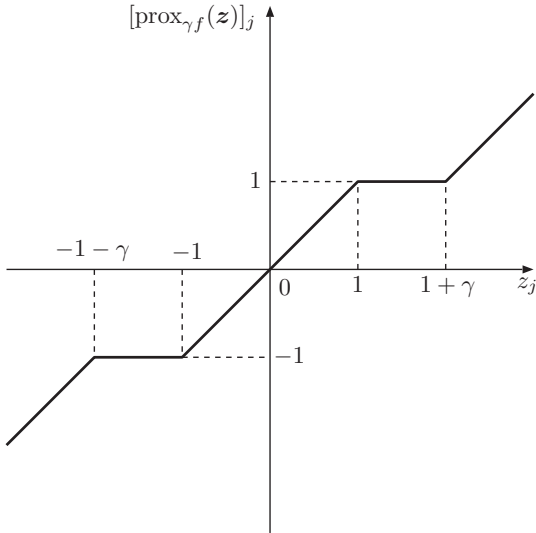


Fig. 1. $[\text{prox}_{\gamma f}(z)]_j$

D. LLR calculation

Although the estimate \hat{s} can be obtained with the SOAV optimization, we consider to further improve the performance of the proposed scheme by extending the SOAV optimization into the W-SOAV optimization. For the preparation of the extension, we here consider to approximate the posterior LLR of s_j defined as

$$\Lambda_j := \log \frac{p(s_j = +1 | \mathbf{y})}{p(s_j = -1 | \mathbf{y})} \quad (14)$$

$$= \log \frac{p(\mathbf{y} | s_j = +1)}{p(\mathbf{y} | s_j = -1)}, \quad (15)$$

by using the current estimate \hat{s} . To reduce the computational complexity, we firstly approximate Λ_j as

$$\Lambda_j \approx \log \frac{\prod_{i=1}^{2m} p(y_i | s_j = +1)}{\prod_{i=1}^{2m} p(y_i | s_j = -1)} \quad (16)$$

$$= \sum_{i=1}^{2m} \log \frac{p(y_i | s_j = +1)}{p(y_i | s_j = -1)} \quad (17)$$

by assuming that the observations y_1, \dots, y_{2m} are independent, which means $p(\mathbf{y} | s_j = +1) = \prod_{i=1}^{2m} p(y_i | s_j = +1)$ and $p(\mathbf{y} | s_j = -1) = \prod_{i=1}^{2m} p(y_i | s_j = -1)$. By using the similar idea to the Gaussian approximation in the BP-based detection [7], we rewrite y_i as

$$y_i = h_{i,j} s_j + \sum_{\substack{k=1 \\ k \neq j}}^{2n} h_{i,k} s_k + v_i = h_{i,j} s_j + \xi_i^j \quad (18)$$

where $\xi_i^j = \sum_{k=1, k \neq j}^{2n} h_{i,k} s_k + v_i$. Since ξ_i^j is the sum of $2n - 1$ independent random variables and Gaussian noise, we can approximate it as a Gaussian random variable from the central limit theorem when n is large. We thus calculate (17)

as

$$\sum_{i=1}^{2m} \log \frac{p(y_i | s_j = +1)}{p(y_i | s_j = -1)} \approx \sum_{i=1}^{2m} \frac{2h_{i,j} (y_i - \mu_{\xi_i^j})}{\sigma_{\xi_i^j}^2}, \quad (19)$$

where $\mu_{\xi_i^j}$ and $\sigma_{\xi_i^j}^2$ represent the mean and the variance of ξ_i^j , respectively, which are given by

$$\mu_{\xi_i^j} = \sum_{\substack{k=1 \\ k \neq j}}^{2n} h_{i,k} \mathbb{E}[s_k], \quad (20)$$

$$\sigma_{\xi_i^j}^2 = \sum_{\substack{k=1 \\ k \neq j}}^{2n} h_{i,k}^2 (1 - \mathbb{E}[s_k]^2) + \frac{\sigma_v^2}{2}. \quad (21)$$

Since $\mathbb{E}[s_k]$ is not available in general, we approximate $\mu_{\xi_i^j}$ and $\sigma_{\xi_i^j}^2$ using the current estimates $\hat{s}_1, \dots, \hat{s}_{2n}$ as

$$\mu_{\xi_i^j} \approx \hat{\mu}_{\xi_i^j} := \sum_{\substack{k=1 \\ k \neq j}}^{2n} h_{i,k} \hat{s}'_k, \quad (22)$$

$$\sigma_{\xi_i^j}^2 \approx \hat{\sigma}_{\xi_i^j}^2 := \sum_{\substack{k=1 \\ k \neq j}}^{2n} h_{i,k}^2 (1 - \hat{s}'_k{}^2) + \frac{\sigma_v^2}{2}, \quad (23)$$

where

$$\hat{s}'_j = \begin{cases} -1 & (\hat{s}_j < -1) \\ \hat{s}_j & (-1 \leq \hat{s}_j < 1) \\ 1 & (1 \leq \hat{s}_j) \end{cases} \quad (24)$$

is bounded in $[-1, 1]$ so that $1 - \hat{s}'_k{}^2$ in (23) is not negative. From (17), (19), (22) and (23), the posterior LLR of s_j can be approximated as

$$\Lambda_j \approx \hat{\Lambda}_j := \sum_{i=1}^{2m} \frac{2h_{i,j} \left(y_i - \sum_{\substack{k=1 \\ k \neq j}}^{2n} h_{i,k} \hat{s}'_k \right)}{\sum_{\substack{k=1 \\ k \neq j}}^{2n} h_{i,k}^2 (1 - \hat{s}'_k{}^2) + \frac{\sigma_v^2}{2}}. \quad (25)$$

Since the computational complexity of (25) is $O(mn)$, the complexity for the direct calculation of $\hat{\Lambda}_1, \dots, \hat{\Lambda}_{2n}$ will be $O(mn^2)$. However, we can reduce the complexity to $O(mn)$ by calculating and storing

$$\hat{\mu}_i := \sum_{k=1}^{2n} h_{i,k} \hat{s}'_k, \quad (26)$$

$$\hat{\sigma}_i^2 := \sum_{k=1}^{2n} h_{i,k}^2 (1 - \hat{s}'_k{}^2) + \frac{\sigma_v^2}{2}, \quad (27)$$

in advance. Since (26) and (27) can be calculated with the complexity of $O(n)$, we can obtain all of $\hat{\mu}_i$ and $\hat{\sigma}_i^2$ ($i = 1, \dots, 2m$) with $O(mn)$. By using $\hat{\mu}_i$ and $\hat{\sigma}_i^2$, (22) and (23) are rewritten as

$$\hat{\mu}_{\xi_i^j} = \hat{\mu}_i - h_{i,j} \hat{s}'_j, \quad (28)$$

$$\hat{\sigma}_{\xi_i^j}^2 = \hat{\sigma}_i^2 - h_{i,j}^2 (1 - \hat{s}'_j{}^2), \quad (29)$$

which can be obtained with $O(1)$. With (26)–(29), (25) can be rewritten as

$$\hat{\Lambda}_j = \sum_{i=1}^{2m} \frac{2h_{i,j} \{y_i - (\hat{\mu}_i - h_{i,j} \hat{s}'_j)\}}{\hat{\sigma}_i^2 - h_{i,j}^2 (1 - \hat{s}'_j{}^2)} \quad (30)$$

and hence the complexity needed for the calculation of each $\hat{\Lambda}_j$ is reduced to just $O(m)$. As a result, we can obtain $\hat{\Lambda}_1, \dots, \hat{\Lambda}_{2n}$ with the complexity of $O(mn)$, while the complexity of the SOAV optimization is $O(n^3)$ as described in the previous section.

E. IW-SOAV

From the approximated posterior LLR $\hat{\Lambda}_j$, we approximate the posterior probabilities as $\Pr(s_j = +1 | \mathbf{y}) \approx w_j^+$ and $\Pr(s_j = -1 | \mathbf{y}) \approx w_j^-$ by using (14), where

$$w_j^+ := \frac{e^{\hat{\Lambda}_j}}{1 + e^{\hat{\Lambda}_j}}, \quad w_j^- := \frac{1}{1 + e^{\hat{\Lambda}_j}}. \quad (31)$$

By using the approximated probabilities as the prior information, we extend the problem (8) to the W-SOAV optimization problem as

$$\hat{\mathbf{s}} = \underset{\mathbf{z} \in \mathbb{R}^{2n}}{\operatorname{argmin}} \left(\sum_{j=1}^{2n} (w_j^+ |z_j - 1| + w_j^- |z_j + 1|) + \frac{\alpha}{2} \|\mathbf{y} - \mathbf{H}\mathbf{z}\|_2^2 \right). \quad (32)$$

If there is no prior information about \mathbf{s} , i.e., $w_j^+ = w_j^- = 1/2$, the optimization problem (32) is equivalent to (8). If $w_j^+ > w_j^-$, then $\underset{z_j \in \mathbb{R}}{\operatorname{argmin}} (w_j^+ |z_j - 1| + w_j^- |z_j + 1|) = 1$ and hence the solution of z_j in (32) tends to take the value close to 1, and vice versa. The optimization problem (32) can also be solved by using the Douglas-Rachford algorithm. The proximity operator of

$$\gamma f_w(\mathbf{z}) := \gamma \sum_{j=1}^{2n} (w_j^+ |z_j - 1| + w_j^- |z_j + 1|) \quad (33)$$

can be written as

$$[\operatorname{prox}_{\gamma f_w}(\mathbf{z})]_j = \begin{cases} z_j + \gamma & (z_j < -1 - \gamma) \\ -1 & (-1 - \gamma \leq z_j < -1 - d_j \gamma) \\ z_j + d_j \gamma & (-1 - d_j \gamma \leq z_j < 1 - d_j \gamma) \\ 1 & (1 - d_j \gamma \leq z_j < 1 + \gamma) \\ z_j - \gamma & (1 + \gamma \leq z_j) \end{cases}, \quad (34)$$

where $d_j = w_j^+ - w_j^-$. Figure 2 shows an example of $[\operatorname{prox}_{\gamma f_w}(\mathbf{z})]_j$ for the case with $w_j^+ > w_j^-$. By solving the optimization problem (32) via the Douglas-Rachford algorithm with $\operatorname{prox}_{\gamma f_w}$ and $\operatorname{prox}_{\gamma g}$, a new estimate of the transmitted signal vector \mathbf{s} can be obtained.

To implement the idea of W-SOAV, we propose an iterative approach, referred to as IW-SOAV. In each iteration of IW-SOAV, we use the estimate obtained in the previous iteration as the prior information. By calculating the weights w_j^+ and

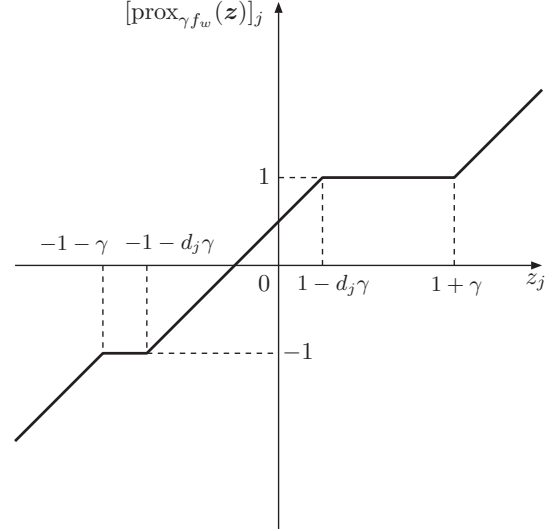


Fig. 2. An example of $[\operatorname{prox}_{\gamma f_w}(\mathbf{z})]_j$ ($w_j^+ > w_j^-$)

w_j^- from the estimate and solving the W-SOAV optimization problem, we can obtain an improved estimate of \mathbf{s} . The proposed algorithm of IW-SOAV is summarized as follows:

Algorithm 2. (Proposed signal detection via IW-SOAV)

- 1) Let $\hat{\mathbf{s}} = \mathbf{0}$ and iterate a)–d) for L times.
 - a) Calculate $\hat{\Lambda}_j$ with (30).
 - b) Compute w_j^+ and w_j^- with (31).
 - c) Fix $\varepsilon \in (0, 1)$, $\gamma > 0$, and $\mathbf{r}_0 \in \mathbb{R}^{2n}$, and set the number of iterations in the W-SOAV optimization K_{itr} .
 - d) For $k = 0, 1, 2, \dots, K_{\text{itr}}$, iterate

$$\begin{cases} \mathbf{z}_k = \operatorname{prox}_{\gamma f_w}(\mathbf{r}_k) \\ \theta_k \in [\varepsilon, 2 - \varepsilon] \\ \mathbf{r}_{k+1} = \mathbf{r}_k + \theta_k (\operatorname{prox}_{\gamma g}(2\mathbf{z}_k - \mathbf{r}_k) - \mathbf{z}_k) \end{cases}$$

and let $\hat{\mathbf{s}} = \mathbf{z}_{K_{\text{itr}}}$.

- 2) Obtain $\operatorname{sgn}(\hat{\mathbf{s}})$ as the final estimate of \mathbf{s} .

The computational complexity of IW-SOAV is the same order as that of the Douglas-Rachford algorithm for the SOAV optimization because it is dominated by the matrix inversion $(\mathbf{I} + \alpha\gamma\mathbf{H}^T\mathbf{H})^{-1}$ in (13).

F. Comments for other modulation schemes

Although we assume QPSK modulation throughout this paper, we can extend our proposed scheme for rectangular quadratic amplitude modulation (QAM) symbols. For example, when we use 16-QAM symbols whose real and imaginary parts take $+3, +1, -1, \text{ or } -3$, the corresponding SOAV optimization problem can be obtained by replacing $\frac{1}{2}\|\mathbf{z} - \mathbf{1}\|_1 + \frac{1}{2}\|\mathbf{z} + \mathbf{1}\|_1$ in (8) with $f_{16\text{-QAM}}(\mathbf{z}) = \frac{1}{4}\|\mathbf{z} - 3 \cdot \mathbf{1}\|_1 + \frac{1}{4}\|\mathbf{z} - \mathbf{1}\|_1 + \frac{1}{4}\|\mathbf{z} + \mathbf{1}\|_1 + \frac{1}{4}\|\mathbf{z} + 3 \cdot \mathbf{1}\|_1$. The optimization problem for 16-QAM symbols can also be solved via the Douglas-Rachford algorithm with the proximity operator of $\gamma f_{16\text{-QAM}}$, which can be obtained by the direct calculation of the definition. Although the proximity operator is a bit more complicated than (12), the computational

complexity to estimate the transmitted signal vector is almost the same as that for the case with QPSK modulation. We can also extend the W-SOAV optimization for such case by using the LLR calculation for 16-QAM symbols [28]. However, for some modulation methods such as 8-phase shift keying (PSK), we cannot directly apply our proposed method. When we use 8-PSK symbols with the alphabet $\left\{1, \frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}, j, -\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}, -1, -\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}, -j, \frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}}\right\}$, the proposed method may provide inappropriate estimates, such as $1 - j\frac{1}{\sqrt{2}}$, 0, and $\frac{1}{\sqrt{2}}$.

III. PROPOSED JOINT DETECTION AND DECODING FOR CODED MIMO SYSTEMS

Since the proposed W-SOAV optimization can use the posterior LLRs of transmitted symbols as the prior information, we can integrate it with soft channel decoding schemes, e.g., LDPC codes or turbo codes. In this section, we propose a joint detection and decoding scheme using the W-SOAV optimization for coded massive overloaded MIMO systems.

Figure 3 shows the system model of the coded MIMO with n transmit antennas and m ($< n$) receive antennas. In the transmitter, Q information bits are encoded into P coded bits by a channel encoder with the code rate $R = Q/P$. For simplicity, P is assumed to be a multiple of $2n$. P coded bits are then modulated into $P/2$ QPSK symbols and sent from n transmit antennas over $T = P/2n$ symbols time.

The received signal vector at time $t \in \{1, \dots, T\}$ is given by

$$\tilde{\mathbf{y}}^{(t)} = \tilde{\mathbf{H}}^{(t)} \tilde{\mathbf{s}}^{(t)} + \tilde{\mathbf{v}}^{(t)}, \quad (35)$$

where $\tilde{\mathbf{H}}^{(t)}$, $\tilde{\mathbf{s}}^{(t)}$, and $\tilde{\mathbf{v}}^{(t)}$ are the channel matrix, the transmitted signal vector, and the noise vector at time t , respectively. We can convert (35) into the real signal model

$$\mathbf{y}^{(t)} = \mathbf{H}^{(t)} \mathbf{s}^{(t)} + \mathbf{v}^{(t)} \quad (36)$$

in the same manner as the transformation from (2) to (4). In the proposed detection and decoding, we iteratively perform the detection with the W-SOAV optimization and the channel decoding to update LLRs of transmitted symbols. The detector obtains the estimate $\hat{\mathbf{s}}^{(t)}$ of $\mathbf{s}^{(t)}$ with the W-SOAV optimization using the information from the channel decoder except for the first iteration. Specifically, by using the posterior LLR obtained at the output of the channel decoder as

$$\lambda_j^{(t)} = \log \frac{p(s_j^{(t)} = +1 | \mathbf{y}^{(t)})}{p(s_j^{(t)} = -1 | \mathbf{y}^{(t)})}, \quad (37)$$

the weight parameters of the W-SOAV optimization are given by

$$w_j^+(t) = \frac{e^{\lambda_j^{(t)}}}{1 + e^{\lambda_j^{(t)}}}, \quad w_j^-(t) = \frac{1}{1 + e^{\lambda_j^{(t)}}} \quad (38)$$

as in (31). After the detection via the W-SOAV optimization with the above $w_j^+(t)$ and $w_j^-(t)$, we calculate the posterior LLRs $\hat{\Lambda}^{(t)} = [\hat{\Lambda}_1^{(t)} \dots \hat{\Lambda}_{2n}^{(t)}]^T$, where $\hat{\Lambda}_j^{(t)}$ is given by (30). Using all LLRs $\hat{\Lambda}^{(1)}, \dots, \hat{\Lambda}^{(T)}$ as the input, the decoder

performs the soft channel decoding and outputs new posterior LLRs $\lambda^{(1)}, \dots, \lambda^{(T)}$ to the MIMO detector, where $\lambda^{(t)} = [\lambda_1^{(t)} \dots \lambda_{2n}^{(t)}]^T$. After a certain number of the iterations of the detection and decoding, the decoder outputs the decoded bits as the final estimate of the transmitted information bits.

IV. PERFORMANCE ANALYSIS

In this section, we give an upper bound of the ℓ_2 norm of the error vector between the true transmitted signal vector and its estimate obtained by the W-SOAV optimization. Since the SOAV optimization is partly based on the idea of compressed sensing as described in Sec. II, we use RIP considered in the performance analysis for compressed sensing [25].

Definition 1 (K -sparse vector). A vector $\mathbf{x} \in \mathbb{R}^N$ is said to be K -sparse if it has at most K non-zero elements.

Definition 2 (RIP). A matrix Φ satisfies RIP of order K if there is a constant $\delta_K \in (0, 1)$ so that

$$(1 - \delta_K) \|\mathbf{x}\|_2^2 \leq \|\Phi \mathbf{x}\|_2^2 \leq (1 + \delta_K) \|\mathbf{x}\|_2^2 \quad (39)$$

holds for all K -sparse vector \mathbf{x} . The constant δ_K is called K -restricted isometry constant.

The W-SOAV optimization problem (32) is equivalent to

$$\begin{aligned} & \underset{\mathbf{z} \in \mathbb{R}^{2n}}{\text{minimize}} \sum_{j=1}^{2n} (w_j^+ |z_j - 1| + w_j^- |z_j + 1|) \\ & \text{subject to } \|\mathbf{y} - \mathbf{H} \mathbf{z}\|_2 \leq \varepsilon \end{aligned} \quad (40)$$

for a proper $\varepsilon \geq 0$ corresponding to α . For the solution of (40) and the true transmitted signal vector, we have the following theorem:

Theorem 2. Let $\hat{\mathbf{s}}$ be the solution of (40) and δ_{2K} be the $2K$ -restricted isometry constant of \mathbf{H} . Assume that the true transmitted signal vector \mathbf{s} satisfies the constraint in (40), i.e., $\|\mathbf{y} - \mathbf{H} \mathbf{s}\|_2 \leq \varepsilon$, and define a set of indices $\mathcal{W} = \{j \mid s_j = 1, w_j^+ < w_j^-\} \cup \{j \mid s_j = -1, w_j^+ > w_j^-\}$, where $s_j \neq \underset{z_j \in \mathbb{R}}{\text{argmin}} (w_j^+ |z_j - 1| + w_j^- |z_j + 1|)$. If the inequality

$$\delta_{2K} < \frac{\bar{w}}{\sqrt{2} + \bar{w}} \quad (41)$$

holds, then we have

$$\|\hat{\mathbf{s}} - \mathbf{s}\|_2 \leq \tau(\bar{w} - \rho)^{-1}(1 + \bar{w})\varepsilon, \quad (42)$$

where

$$K = |\mathcal{J}|, \quad \bar{w} = \min_{j \in \mathcal{J}^c} |w_j^+ - w_j^-|, \quad (43)$$

$$\tau = \frac{2\sqrt{1 + \delta_{2K}}}{1 - \delta_{2K}}, \quad \rho = \frac{\sqrt{2}\delta_{2K}}{1 - \delta_{2K}}, \quad (44)$$

and $\mathcal{J} \subset \{1, \dots, 2n\}$ is a set of indices satisfying $\mathcal{J} \supset \mathcal{W}$.

Proof: See Appendix A. ■

The procedure of the proof for this theorem is partly similar to that in [25] for compressed sensing and our result is a kind of generalization of that for compressed sensing. The condition (41) depends on δ_{2K} and \bar{w} , while the corresponding condition for compressed sensing $\delta_{2K} < \sqrt{2} - 1$ depends only

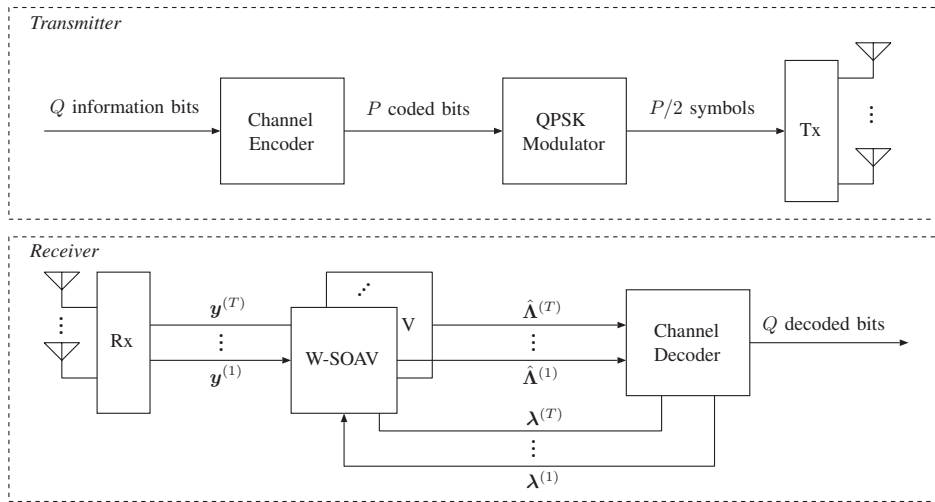


Fig. 3. Model of coded MIMO systems

on δ_{2K} . If \bar{w} takes its maximum value, i.e., $\bar{w} = 1$, then (41) and (42) are equivalent to the result for compressed sensing. Indeed, (41) and (42) for $\bar{w} = 1$ can be written as

$$\delta_{2K} < \frac{1}{\sqrt{2} + 1} = \sqrt{2} - 1, \quad (45)$$

$$\|\hat{\mathbf{s}} - \mathbf{s}\|_2 \leq 2\tau(1 - \rho)^{-1}\varepsilon, \quad (46)$$

which correspond to the result for the reconstruction of a K -sparse vector via the ℓ_1 optimization problem. However, the smaller \bar{w} is, the severer the condition (41) is because $\bar{w}/(\sqrt{2} + \bar{w})$ is the monotonically increasing function of \bar{w} . Since the upper bound in (42) is the monotonically decreasing function of \bar{w} , it has a large value for a small \bar{w} .

In the noise-free case, we can set $\varepsilon = 0$, and (42) shows that the norm of the reconstruction error with the W-SOAV optimization is upper bounded by 0. Namely, Theorem 2 shows the sufficient condition for the exact reconstruction with the W-SOAV optimization in the noise-free case.

One of drawbacks of the performance analysis based on RIP is that it is difficult in general to calculate the restricted isometry constant δ_K for a specific matrix due to the infeasible computational complexity. For random matrices, however, some asymptotical results about RIP have been obtained [29] and hence they might be used with our theorem.

V. SIMULATION RESULTS

In this section, we evaluate the BER performance of the proposed scheme via computer simulations. The parameters of the Douglas-Rachford algorithm are set as $\mathbf{r}_0 = \mathbf{0}$, $\varepsilon = 0.1$, $\gamma = 1$, and $\theta_k = 1.9$ ($k = 0, 1, \dots, K_{\text{itr}}$), which give fast convergence.

A. Uncoded MIMO system

Figures 4–6 show the BER performance for uncoded MIMO systems with $(n, m) = (50, 32)$, $(100, 64)$, and $(150, 96)$, respectively. In the figures, we assume flat Rayleigh fading channels and set $\tilde{\mathbf{H}} = \tilde{\mathbf{H}}_{\text{i.i.d.}}$, which is composed of i.i.d. complex Gaussian random variables with zero mean and unit

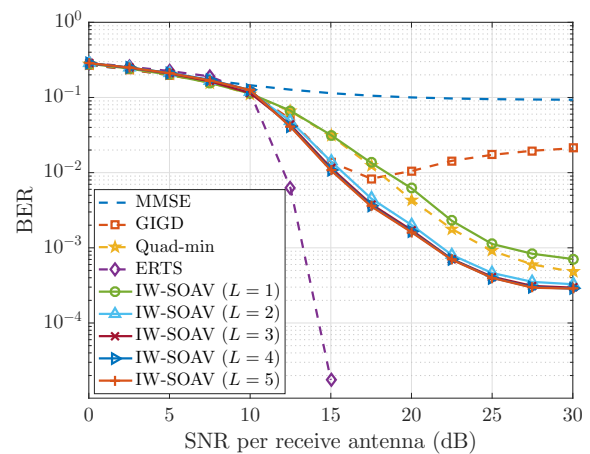


Fig. 4. BER performance for uncoded MIMO with $(n, m) = (50, 32)$

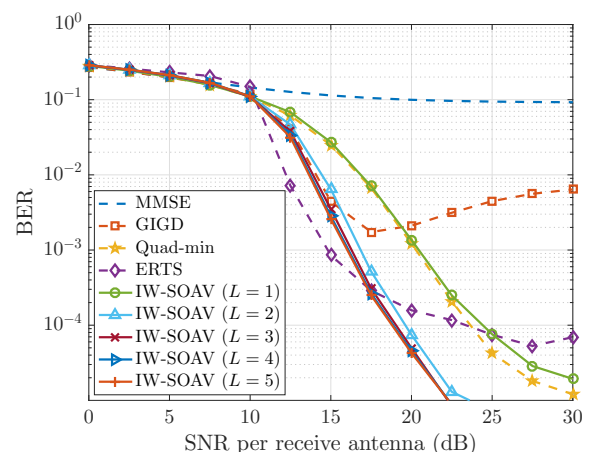


Fig. 5. BER performance for uncoded MIMO with $(n, m) = (100, 64)$

variance. The number of iterations in the Douglas-Rachford algorithm is fixed to $K_{\text{itr}} = 50$, which is sufficiently large for the convergence of the algorithm. The parameter α in (32) is selected as shown in TABLE I, which is determined from simulation results. Note that the ratio $m/n = 0.64$ is identical

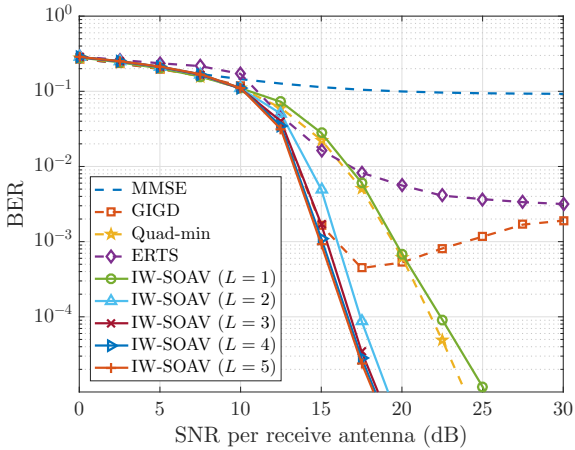


Fig. 6. BER performance for uncoded MIMO with $(n, m) = (150, 96)$

TABLE I
THE VALUE OF α IN (32)

SNR per receive antenna (dB)	0–10	12.5–20	22.5	25–30
α	0.01	0.1	0.3	1

for all the figures. We denote the linear MMSE detection by “MMSE”, the BP-based detection [7] by “GIGD”, the detection with a quadratic programming [20] by “Quad-min”, and the massive overloaded MIMO signal detection proposed in [18] by “ERTS”. For the implementation of “Quad-min”, we use the CVX toolbox [30] for MATLAB. The parameters of ERTS are the same as those in [18], e.g., the maximum number of RTSs is $N_{RTS} = 500$ and the maximum number of iterations in RTS is $N_{itr} = 300$. “IW-SOAV” denotes the proposed scheme shown in Algorithm 2 and L is the number of iterative W-SOAV optimizations in IW-SOAV. In Fig. 4, where $n = 50$, the performance of IW-SOAV is inferior to that of ERTS. In Fig. 5 with $n = 100$, however, the performance of ERTS has degraded and IW-SOAV has better performance in high SNR. In Fig. 6 for further large-scale MIMO, IW-SOAV outperforms the other schemes. The reason for the performance degradation of ERTS is that, if the number of transmit antennas is large, RTS often fails to find the true transmitted signal vector due to the huge number of candidates of the transmitted vector. Although we may get better performance with ERTS by increasing the number of RTSs, the computational complexity could be prohibitive to achieve comparable performance as IW-SOAV. Specifically, the computational complexity of ERTS is given by $O(n^3) + O(N_{RTS}n^2)$ in the worst case, and since the number of all candidates of the transmit signal vector increases exponentially with the number of transmit antennas, the required N_{RTS} to keep good performance will increase more rapidly than n . On the other hand, the computational complexity of IW-SOAV is $O(n^3)$ as described in Sec. II. To compare the complexity, we evaluate the average computation time to detect a transmitted symbol vector versus n and the corresponding BER performance for the fixed ratio $m/n = 2/3$ and the SNR per receive antenna of 17.5 dB in Figs. 7 and 8, respectively. The simulation is conducted by using a computer with 2 GHz Intel Core i7-3667U and 8 GB memory.

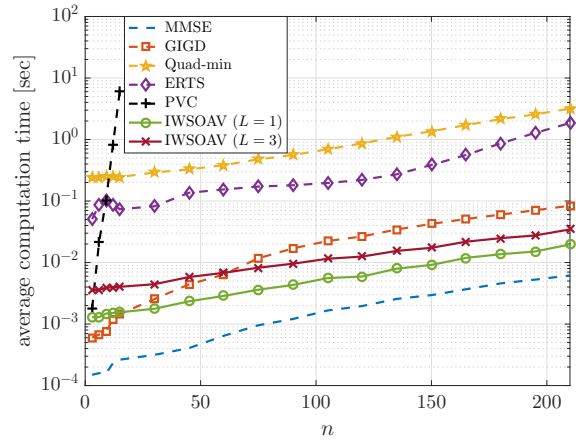


Fig. 7. Average computation time versus n for uncoded MIMO with $m/n = 2/3$ and the SNR per receive antenna of 17.5 dB

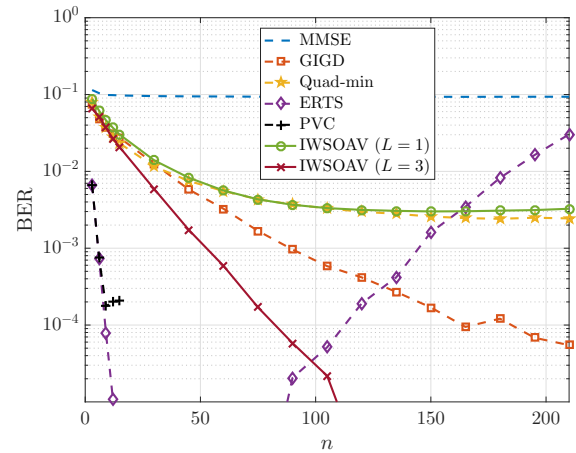


Fig. 8. BER performance versus n for uncoded MIMO with $m/n = 2/3$ and the SNR per receive antenna of 17.5 dB

In the figures, “PVC” represents the signal detection scheme proposed in [15], which is intended for small-scale overloaded MIMO systems. Although PVC can achieve a comparable BER performance to ML detection for small-scale MIMO systems, its average computation time rapidly increases along with n . In Fig. 8, the BER performance of ERTS severely degrades for large n . This is because the maximum number of RTSs is limited as $N_{RTS} = 500$ to avoid the prohibitive computational complexity, while the number of candidates of the transmitted signal vector exponentially increases along with n . Compared to the conventional detection schemes, the proposed IW-SOAV can achieve better BER performance with lower complexity in large-scale overloaded MIMO systems.

Figure 9 shows the BER performance versus the number of receive antennas m for $n = 150$ and the SNR per receive antenna of 20 dB. We can observe that IW-SOAV with $L = 5$ requires less antennas than other schemes to achieve a certain BER performance. For $BER = 10^{-4}$, IW-SOAV can reduce more than ten receive antennas compared to ERTS.

In Fig. 10, we also show the BER performance for spatially correlated MIMO channels with $(n, m) = (100, 64)$ and $\tilde{\mathbf{H}} =$

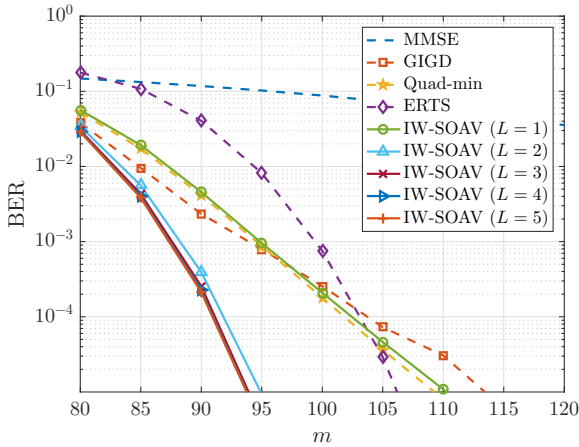


Fig. 9. BER performance versus m for uncoded MIMO with $n = 150$ and the SNR per receive antenna of 20 dB

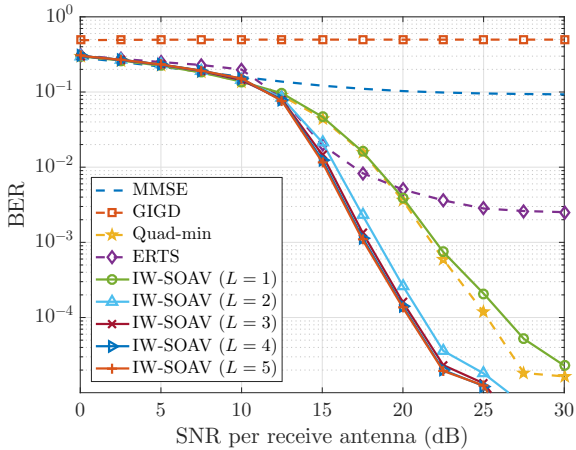


Fig. 10. BER performance for spatially correlated MIMO with $(n, m) = (100, 64)$

$\Phi_R^{\frac{1}{2}} \tilde{H}_{i.i.d.} \Phi_T^{\frac{1}{2}}$, where positive definite matrices $\Phi_R \in \mathbb{C}^{m \times m}$ and $\Phi_T \in \mathbb{C}^{n \times n}$ are called receive and transmit correlation matrices, respectively [31]. We further assume a linear array with equally spaced antennas in both the receiver and the transmitter, and define $[\Phi_R]_{i,j} = J_0(|i-j| \cdot 2\pi d_R/\lambda)$ and $[\Phi_T]_{i,j} = J_0(|i-j| \cdot 2\pi d_T/\lambda)$, where $[\Phi_R]_{i,j}$ and $[\Phi_T]_{i,j}$ denote the (i, j) element of Φ_R and Φ_T , respectively. Here, $J_0(\cdot)$ represents the zeroth-order Bessel function of the first kind and λ denotes the wavelength. d_R and d_T are the antenna spacing at the receiver and the transmitter, respectively, and we set to $d_R = d_T = 0.5\lambda$ in the simulations. From Fig. 10, we can see that the proposed scheme can achieve better performance compared to the conventional schemes even in the spatially correlated MIMO channels, while the performance of GIGD and ERTS is degraded significantly.

B. LDPC coded MIMO system

Figures 11 and 12 show the BER performance of the proposed signal detection and decoding for LDPC coded MIMO with $(n, m) = (100, 64)$. The parameters of the algorithm are set as $K_{\text{itr}} = 30$ and $\alpha = 0.01$. The code

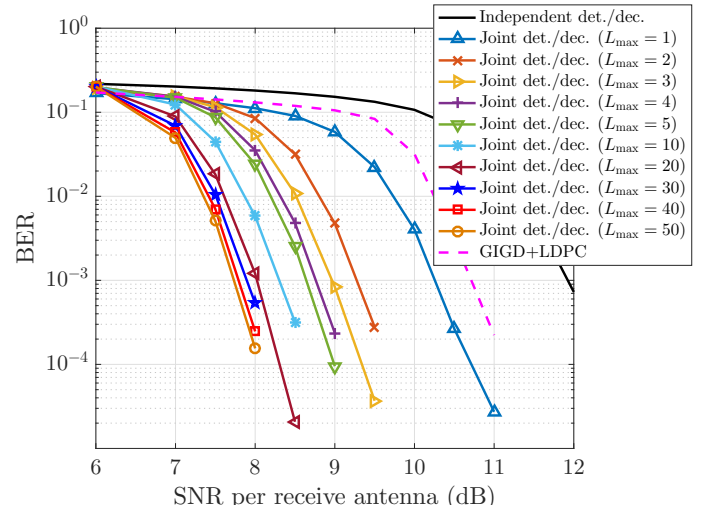


Fig. 11. BER performance for LDPC coded MIMO with $(n, m) = (100, 64)$, $R = 1/2$ and $N = 4000$

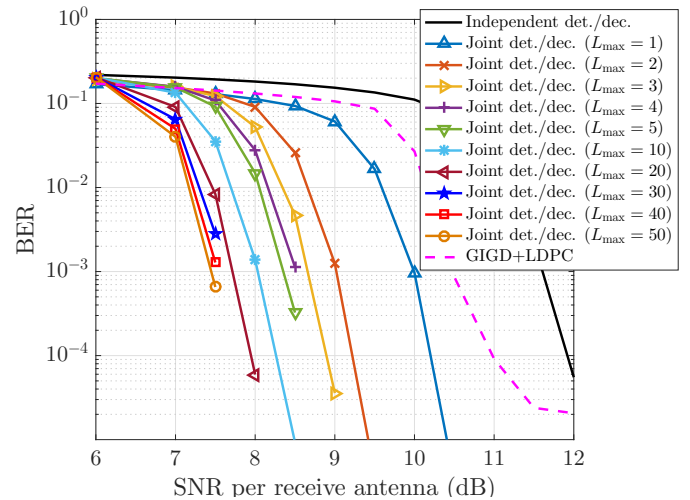


Fig. 12. BER performance for LDPC coded MIMO with $(n, m) = (100, 64)$, $R = 1/2$ and $N = 8000$

rate is $R = 1/2$, and the column and row weights of the parity check matrix are three and six, respectively. In the figures, the code length are $N = 4000$ and 8000 , respectively. We represent the proposed joint detection and decoding by “Joint det./dec.”, where L_{max} indicates the maximum number of iterative W-SOAV optimizations. Even before the L_{max} th iteration, the LDPC decoder outputs the final estimate of the information bits if the decoded bits satisfy all parity check constraints. From the figures, we can see that, as the iteration proceeds, the performance of the joint detection and decoding is considerably improved via LLR update between the W-SOAV optimization and the LDPC decoding. For comparison, we also plot the performance of the independent detection and LDPC decoding (“Independent det./dec.”), where IW-SOAV with $L = 5$ is used as the detection scheme. Moreover, “GIGD+LDPC” shows the performance of the joint detection and decoding with GIGD and LDPC decoding, which are integrated in the same manner as in Fig. 3. The number of

outer iterations between the detector and the decoder is set to 5. We can see that the proposed joint detection and decoding achieves much better performance than the scheme with GIGD and the independent approach. Each element of the estimate $\hat{\mathbf{s}}$ obtained by IW-SOAV with $L = 5$ is almost hard decision, i.e., close to 1 or -1 , and hence LDPC decoding in the independent approach has poor performance even compared to the case with $L = 1$. The figures also show that the performance is improved as the code length increases.

VI. CONCLUSION

In this paper, we have proposed the massive overloaded MIMO signal detection scheme, named IW-SOAV, which iteratively solves the convex W-SOAV optimization problem with updating weights in the cost function. For coded MIMO systems, we have also proposed the joint detection and decoding using the W-SOAV optimization. The proposed joint scheme iteratively updates the LLRs of the transmitted symbols between the detector and the decoder. As the performance analysis of the proposed W-SOAV optimization, we have evaluated the size of the error vector by using RIP. The theoretical analysis can be regarded as a generalization of the results for compressed sensing with the ℓ_1 optimization. Simulation results show that IW-SOAV can achieve much better performance than conventional schemes, especially in large-scale overloaded MIMO systems. Even for spatially correlated MIMO channels, IW-SOAV outperforms conventional schemes. Moreover, the average computation time for IW-SOAV is less than that for other detection schemes intended for massive overloaded MIMO systems. It is also shown for LDPC coded massive overloaded MIMO systems that the proposed joint detection and decoding can achieve better performance compared to the individual detection and decoding. Future work includes an extension of the proposed scheme for multiuser MIMO and the evaluation of its system-level performance.

APPENDIX A PROOF OF THEOREM 2

Let $\mathbf{e} = [e_1 \cdots e_{2n}]^T = \hat{\mathbf{s}} - \mathbf{s}$. \mathcal{T}_1 denotes the set of the indices corresponding to the K largest elements in $\mathbf{e}_{\mathcal{J}}$. Similarly, \mathcal{T}_2 denotes the set of the indices corresponding to the K largest elements in $\mathbf{e}_{(\mathcal{J} \cup \mathcal{T}_1)^c}$. We also define $\mathcal{T}_3, \mathcal{T}_4, \dots$ in the same manner. By definition, the vectors $\mathbf{e}_{\mathcal{J}}, \mathbf{e}_{\mathcal{T}_1}, \mathbf{e}_{\mathcal{T}_2}, \dots$ are K -sparse. Since $\mathcal{J}, \mathcal{T}_1, \mathcal{T}_2, \dots$ are disjoint, $\|\mathbf{e}\|_2$ is bounded as

$$\|\mathbf{e}\|_2 \leq \|\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2 + \sum_{u \geq 2} \|\mathbf{e}_{\mathcal{T}_u}\|_2. \quad (47)$$

We firstly evaluate $\|\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2$. By using the fact that $\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}$

is $2K$ -sparse, we have

$$(1 - \delta_{2K}) \|\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2^2 \leq \|\mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2^2 \quad (48)$$

$$= \langle \mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}, \mathbf{H}\mathbf{e} \rangle - \left\langle \mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}, \sum_{u \geq 2} \mathbf{H}\mathbf{e}_{\mathcal{T}_u} \right\rangle \quad (49)$$

$$\leq |\langle \mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}, \mathbf{H}\mathbf{e} \rangle| + \left| \left\langle \mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}, \sum_{u \geq 2} \mathbf{H}\mathbf{e}_{\mathcal{T}_u} \right\rangle \right|. \quad (50)$$

The first term in (50) is bounded as

$$|\langle \mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}, \mathbf{H}\mathbf{e} \rangle| \leq \|\mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2 \|\mathbf{H}\mathbf{e}\|_2. \quad (51)$$

Using the inequalities of

$$\|\mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2^2 \leq (1 + \delta_{2K}) \|\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2^2 \quad (52)$$

and

$$\|\mathbf{H}\mathbf{e}\|_2 = \|\mathbf{H}(\hat{\mathbf{s}} - \mathbf{s})\|_2 \quad (53)$$

$$= \|(\mathbf{H}\hat{\mathbf{s}} - \mathbf{y}) - (\mathbf{H}\mathbf{s} - \mathbf{y})\|_2 \quad (54)$$

$$\leq \|\mathbf{H}\hat{\mathbf{s}} - \mathbf{y}\|_2 + \|\mathbf{H}\mathbf{s} - \mathbf{y}\|_2 \quad (55)$$

$$\leq 2\varepsilon, \quad (56)$$

we obtain

$$|\langle \mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}, \mathbf{H}\mathbf{e} \rangle| \leq 2\varepsilon \sqrt{1 + \delta_{2K}} \|\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2. \quad (57)$$

The second term in (50) is bounded as

$$\left| \left\langle \mathbf{H}\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}, \sum_{u \geq 2} \mathbf{H}\mathbf{e}_{\mathcal{T}_u} \right\rangle \right| \leq \sum_{u \geq 2} (|\langle \mathbf{H}\mathbf{e}_{\mathcal{J}}, \mathbf{H}\mathbf{e}_{\mathcal{T}_u} \rangle| + |\langle \mathbf{H}\mathbf{e}_{\mathcal{T}_1}, \mathbf{H}\mathbf{e}_{\mathcal{T}_u} \rangle|) \quad (58)$$

$$\leq \sum_{u \geq 2} (\delta_{2K} \|\mathbf{e}_{\mathcal{J}}\|_2 \|\mathbf{e}_{\mathcal{T}_u}\|_2 + \delta_{2K} \|\mathbf{e}_{\mathcal{T}_1}\|_2 \|\mathbf{e}_{\mathcal{T}_u}\|_2) \quad (59)$$

$$= \delta_{2K} (\|\mathbf{e}_{\mathcal{J}}\|_2 + \|\mathbf{e}_{\mathcal{T}_1}\|_2) \sum_{u \geq 2} \|\mathbf{e}_{\mathcal{T}_u}\|_2 \quad (60)$$

$$\leq \sqrt{2} \delta_{2K} \|\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2 \sum_{u \geq 2} \|\mathbf{e}_{\mathcal{T}_u}\|_2, \quad (61)$$

where the following lemma [25] is used for the transformation from (58) to (59):

Lemma 1. For two vectors $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^{2n}$ satisfying $|\text{supp}(\mathbf{z}_1)| = K_1, |\text{supp}(\mathbf{z}_2)| = K_2$, and $\text{supp}(\mathbf{z}_1) \cap \text{supp}(\mathbf{z}_2) = \emptyset$, we have

$$|\langle \mathbf{H}\mathbf{z}_1, \mathbf{H}\mathbf{z}_2 \rangle| \leq \delta_{K_1+K_2} \|\mathbf{z}_1\|_2 \|\mathbf{z}_2\|_2, \quad (62)$$

where $\text{supp}(\cdot)$ denotes the set of indices corresponding to nonzero elements.

From (50), (57), and (61), we have

$$(1 - \delta_{2K}) \|\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2^2 \leq 2\varepsilon \sqrt{1 + \delta_{2K}} \|\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2 + \sqrt{2} \delta_{2K} \|\mathbf{e}_{\mathcal{J} \cup \mathcal{T}_1}\|_2 \sum_{u \geq 2} \|\mathbf{e}_{\mathcal{T}_u}\|_2. \quad (63)$$

By dividing both sides with $(1 - \delta_{2K})\|e_{\mathcal{J} \cup \mathcal{T}_1}\|_2$, it follows from (63) that

$$\|e_{\mathcal{J} \cup \mathcal{T}_1}\|_2 \leq \frac{2\sqrt{1 + \delta_{2K}}}{1 - \delta_{2K}}\varepsilon + \frac{\sqrt{2}\delta_{2K}}{1 - \delta_{2K}} \sum_{u \geq 2} \|e_{\mathcal{T}_u}\|_2 \quad (64)$$

$$= \tau\varepsilon + \rho \sum_{u \geq 2} \|e_{\mathcal{T}_u}\|_2. \quad (65)$$

Next, we evaluate $\sum_{u \geq 2} \|e_{\mathcal{T}_u}\|_2$, which appears in (47) and (65). For $u \geq 2$, we have

$$\|e_{\mathcal{T}_u}\|_2 \leq \sqrt{K}\|e_{\mathcal{T}_u}\|_\infty \leq \frac{1}{\sqrt{K}}\|e_{\mathcal{T}_{u-1}}\|_1, \quad (66)$$

and hence

$$\sum_{u \geq 2} \|e_{\mathcal{T}_u}\|_2 \leq \sum_{u \geq 2} \frac{1}{\sqrt{K}}\|e_{\mathcal{T}_{u-1}}\|_1 \quad (67)$$

$$= \frac{1}{\sqrt{K}} \sum_{u \geq 1} \|e_{\mathcal{T}_u}\|_1 \quad (68)$$

$$= \frac{1}{\sqrt{K}}\|e_{\mathcal{J}^c}\|_1. \quad (69)$$

To evaluate $\|e_{\mathcal{J}^c}\|_1$, we use the fact that $\hat{s} = s + e$ is the solution of the optimization problem (40). Since $\hat{s} = s + e$ is the minimizer of the objective function, we have

$$\begin{aligned} & \sum_{j=1}^{2n} (w_j^+ |s_j - 1| + w_j^- |s_j + 1|) \\ & \geq \sum_{j=1}^{2n} (w_j^+ |s_j + e_j - 1| + w_j^- |s_j + e_j + 1|) \quad (70) \end{aligned}$$

$$\begin{aligned} & = \sum_{j \in \mathcal{J}} (w_j^+ |s_j + e_j - 1| + w_j^- |s_j + e_j + 1|) \\ & \quad + \sum_{j \in \mathcal{J}^c} (w_j^+ |s_j + e_j - 1| + w_j^- |s_j + e_j + 1|) \quad (71) \end{aligned}$$

$$\begin{aligned} & \geq \sum_{j \in \mathcal{J}} \{w_j^+ (|s_j - 1| - |e_j|) + w_j^- (|s_j + 1| - |e_j|)\} \\ & \quad + \sum_{j \in \mathcal{J}^c} (w_j^+ |s_j + e_j - 1| + w_j^- |s_j + e_j + 1|), \quad (72) \end{aligned}$$

which results in

$$\begin{aligned} & \sum_{j \in \mathcal{J}^c} (w_j^+ |s_j - 1| + w_j^- |s_j + 1|) + \|e_{\mathcal{J}}\|_1 \\ & \geq \sum_{j \in \mathcal{J}^c} (w_j^+ |s_j + e_j - 1| + w_j^- |s_j + e_j + 1|). \quad (73) \end{aligned}$$

We split \mathcal{J}^c into $\mathcal{J}^c \cap \mathcal{S}^+$ and $\mathcal{J}^c \cap \mathcal{S}^-$, where $\mathcal{S}^+ = \{j \mid s_j = 1\}$ and $\mathcal{S}^- = \{j \mid s_j = -1\}$. The inequality (73) can be rewritten as

$$\begin{aligned} & \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^+} (w_j^+ |s_j - 1| + w_j^- |s_j + 1|) \\ & \quad + \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^-} (w_j^+ |s_j - 1| + w_j^- |s_j + 1|) + \|e_{\mathcal{J}}\|_1 \\ & \geq \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^+} (w_j^+ |s_j + e_j - 1| + w_j^- |s_j + e_j + 1|) \\ & \quad + \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^-} (w_j^+ |s_j + e_j - 1| + w_j^- |s_j + e_j + 1|), \quad (74) \end{aligned}$$

which gives

$$\begin{aligned} & \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^+} 2w_j^- + \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^-} 2w_j^+ + \|e_{\mathcal{J}}\|_1 \\ & \geq \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^+} (w_j^+ |e_j| + w_j^- |e_j + 2|) \\ & \quad + \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^-} (w_j^+ |e_j - 2| + w_j^- |e_j|) \quad (75) \end{aligned}$$

$$\begin{aligned} & \geq \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^+} \{w_j^+ |e_j| + w_j^- (-|e_j| + 2)\} \\ & \quad + \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^-} \{w_j^+ (-|e_j| + 2) + w_j^- |e_j|\} \quad (76) \end{aligned}$$

$$\begin{aligned} & \geq \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^+} (w_j^+ - w_j^-) |e_j| + \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^-} (w_j^- - w_j^+) |e_j| \\ & \quad + \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^+} 2w_j^- + \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^-} 2w_j^+. \quad (77) \end{aligned}$$

It follows from (77) that

$$\begin{aligned} & \|e_{\mathcal{J}}\|_1 \\ & \geq \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^+} (w_j^+ - w_j^-) |e_j| + \sum_{j \in \mathcal{J}^c \cap \mathcal{S}^-} (w_j^- - w_j^+) |e_j| \quad (78) \end{aligned}$$

$$\geq \bar{w} \|e_{\mathcal{J}^c}\|_1. \quad (79)$$

From (79) and the Cauchy-Schwarz inequality $\|e_{\mathcal{J}}\|_1 / \sqrt{K} \leq \|e_{\mathcal{J}}\|_2$, $\|e_{\mathcal{J}^c}\|_1$ is bounded as

$$\|e_{\mathcal{J}^c}\|_1 \leq \frac{1}{\bar{w}} \|e_{\mathcal{J}}\|_1 \leq \frac{\sqrt{K}}{\bar{w}} \|e_{\mathcal{J}}\|_2 \leq \frac{\sqrt{K}}{\bar{w}} \|e_{\mathcal{J} \cup \mathcal{T}_1}\|_2. \quad (80)$$

Substituting (80) into (69) gives

$$\sum_{u \geq 2} \|e_{\mathcal{T}_u}\|_2 \leq \frac{1}{\bar{w}} \|e_{\mathcal{J} \cup \mathcal{T}_1}\|_2. \quad (81)$$

From (65) and (81), we have

$$\|e_{\mathcal{J} \cup \mathcal{T}_1}\|_2 \leq \tau\varepsilon + \frac{\rho}{\bar{w}} \|e_{\mathcal{J} \cup \mathcal{T}_1}\|_2. \quad (82)$$

Moreover, if $\bar{w} - \rho > 0$, i.e., $\delta_{2K} < \bar{w} / (\sqrt{2} + \bar{w})$, then (82) can be rewritten as

$$\|e_{\mathcal{J} \cup \mathcal{T}_1}\|_2 \leq (\bar{w} - \rho)^{-1} \bar{w} \tau \varepsilon, \quad (83)$$

and thus we have

$$\sum_{u \geq 2} \|e_{\mathcal{T}_u}\|_2 \leq (\bar{w} - \rho)^{-1} \tau \varepsilon. \quad (84)$$

We conclude from (47), (83), and (84) that

$$\|e\|_2 \leq \tau(\bar{w} - \rho)^{-1} (1 + \bar{w}) \varepsilon. \quad (85)$$

ACKNOWLEDGMENT

This work was supported in part by the Grants-in-Aid for Scientific Research no. 15K06064 and 15H02252 from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and the Grant-in-Aid for JSPS Research Fellow no. 17J07055 from Japan Society for the Promotion of Science.

REFERENCES

- [1] A. Chockalingam and B. S. Rajan, *Large MIMO systems*, Cambridge University Press, 2014.
- [2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
- [3] K. V. Vardhan, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity detector for large MIMO systems and multicarrier CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 473–485, Apr. 2008.
- [4] P. Li and R. D. Murch, "Multiple output selection-LAS algorithm in large MIMO systems," *IEEE Commun. Lett.*, vol. 14, pp. 399–401, May 2010.
- [5] N. Srinidhi, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "Near-ML signal detection in large-dimension linear vector channels using reactive tabu search," arXiv:0911.4640v1 [cs.IT] 24 Nov. 2009.
- [6] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Random-restart reactive tabu search algorithm for detection in large-MIMO systems," *IEEE Commun. Lett.*, vol. 14, no. 12, pp. 1107–1109, Dec. 2010.
- [7] T. Wo and P. A. Hoeher, "A simple iterative Gaussian detector for severely delay-spread MIMO channels," in *Proc. IEEE ICC 2007*, pp. 4598–4563, Jun. 2007.
- [8] M. Sellathurai and S. Haykin, "TURBO-BLAST for wireless communications: theory and experiments," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2538–2546, Oct. 2002.
- [9] R. Gallager, "Low-density parity-check codes," *IRE Trans. Inf. Theory*, vol. 8, no. 1, pp. 21–28, Jan. 1962.
- [10] D. J. C. MacKay and R. M. Neal, "Near Shannon limit performance of low density parity check codes," *Electron. Lett.*, vol. 32, no. 18, pp. 1645–1646, Aug. 1996.
- [11] D. J. C. MacKay, David MacKay's Gallager code resources, <http://www.inference.phy.cam.ac.uk/mackay/CodesFiles.html>. (accessed 11 October 2016)
- [12] T. L. Narasimhan, A. Chockalingam, and B. S. Rajan, "Factor graph based joint detection/decoding for LDPC coded large-MIMO systems," in *Proc. VTC Spring 2012*, pp. 1–5, May 2012.
- [13] K. K. Wong and A. Paulraj, "Efficient high-performance decoding for overloaded MIMO antenna systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1833–1843, May 2007.
- [14] M. Pohst, "On the computation of lattice vectors of minimal length, successive minima and reduced basis with applications," *ACM SIGSAM Bulletin*, vol. 15, pp. 37–44, Feb. 1981.
- [15] L. Bai, C. Chen, and J. Choi, "Lattice reduction aided detection for underdetermined MIMO systems: a pre-voting cancellation approach," in *Proc. VTC Spring 2010*, pp. 1–5, May 2010.
- [16] R. Hayakawa, K. Hayashi, and M. Kaneko, "An overloaded MIMO signal detection scheme with slab decoding and lattice reduction," in *Proc. APCC 2015*, pp. 42–46, Oct. 2015.
- [17] R. Hayakawa, K. Hayashi, and M. Kaneko, "Lattice reduction-aided detection for overloaded MIMO using slab decoding," *IEICE Trans. Commun.*, vol. E99-B, no. 8, pp. 1697–1705, Aug. 2016.
- [18] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Low-complexity near-optimal signal detection in underdetermined large-MIMO systems," in *Proc. NCC 2012*, pp. 1–5, Feb. 2012.
- [19] Y. Fadlallah, A. Aïssa-El-Bey, K. Amis, D. Pastor, and R. Pyndiah, "New decoding strategy for underdetermined MIMO transmission using sparse decomposition," in *Proc. EUSIPCO 2013*, pp. 1–5, Sept. 2013.
- [20] Y. Fadlallah, A. Aïssa-El-Bey, K. Amis, D. Pastor, and R. Pyndiah, "New iterative detector of MIMO transmission using sparse decomposition," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3458–3464, Aug. 2015.
- [21] R. Hayakawa, K. Hayashi, H. Sasahara, and M. Nagahara, "Massive overloaded MIMO signal detection via convex optimization with proximal splitting," in *Proc. EUSIPCO 2016*, pp. 1383–1387, Aug.–Sept. 2016.
- [22] M. Nagahara, "Discrete signal reconstruction by sum of absolute values," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1575–1579, Oct. 2015.
- [23] A. Aïssa-El-Bey, D. Pastor, S. M. A. Sbaï, and Y. Fadlallah, "Sparsity-based recovery of finite alphabet solutions to underdetermined linear systems," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 2008–2018, Apr. 2015.
- [24] P. Combettes and J. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, ser. Springer Optimization and Its Applications. Springer New York, vol. 49, pp. 185–212, 2011.
- [25] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *C. R. Acad. Sci.*, ser. I, no. 346, pp. 589–592, 2008.
- [26] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [27] 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, Mar. 2013.
- [28] S. Allpress, C. Luschi, and S. Felix, "Exact and approximated expressions of the log-likelihood ratio for 16-QAM signals," in *Proc. 38th Asilomar Conference on Signal, System and Computer*, Nov. 2004.
- [29] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, Jan. 2008.
- [30] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx>, Dec. 2016.
- [31] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

PLACE
PHOTO
HERE

Ryo Hayakawa received the Bachelor's degree in engineering and the Master's degree in informatics from Kyoto University, Kyoto, Japan, in 2015 and 2017, respectively. He is currently pursuing a Ph.D. degree in the Department of Systems Science, Graduate School of Informatics, Kyoto University. He is a Research Fellow (DC1) of the Japan Society for the Promotion of Science (JSPS). His research interests include signal processing and wireless communication. He is a student member of IEICE.

PLACE
PHOTO
HERE

Kazunori Hayashi is currently a Professor at Graduate School of Engineering, Osaka City 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, and an Associate Professor from 2007 to 2017 at Graduate School of Informatics, Kyoto 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, and the IEICE Communications Society Best Tutorial Paper Award in 2013. He is a member of IEICE, APSIPA, and ISIC.