PAPER *Special Section on Information Theory and Its Applications*

Error Recovery for Massive MIMO Signal Detection via Reconstruction of Discrete-Valued Sparse Vector*

Ryo HAYAKAWA†a), ***Student Member and*** **Kazunori HAYASHI**††b), ***Member***

SUMMARY In this paper, we propose a novel error recovery method for massive multiple-input multiple-output (MIMO) signal detection, which improves an estimate of transmitted signals by taking advantage of the sparsity and the discreteness of the error signal. We firstly formulate the error recovery problem as the maximum a posteriori (MAP) estimation and then relax the MAP estimation into a convex optimization problem, which reconstructs a discrete-valued sparse vector from its linear measurements. By using the restricted isometry property (RIP), we also provide a theoretical upper bound of the size of the reconstruction error with the optimization problem. Simulation results show that the proposed error recovery method has better bit error rate (BER) performance than that of the conventional error recovery method.

key words: *massive MIMO, signal detection, sum-of-absolute-value optimization, restricted isometry property*

1. Introduction

Because of the significant increase of the required data rate and throughput in wireless communications systems, much attention has been paid to massive multiple-input multiple-output (MIMO) systems with tens or hundreds of antennas [1]. For massive MIMO systems, a low-complexity signal detection scheme will be required because the computational complexity increases along with the number of antennas. Although linear signal detections, such as the zero forcing (ZF) and the minimum mean-square-error (MMSE) detections, can be possible candidates for massive MIMO systems, the performance is much inferior to that of the optimal maximum likelihood (ML) detection. To achieve nearly optimal performance, some non-linear detection schemes have also been proposed. The likelihood ascent search (LAS) [2] and the reactive tabu search (RTS) [3] are non-linear detection schemes based on the local neighborhood search of likelihood. The belief propagation-based detection [4] and the convex optimization-based detection [5] have also been proposed.

As an another approach for massive MIMO signal detection, post-detection sparse error recovery (PDSR) has been proposed [6]. It improves the estimate obtained by some conventional detection methods, such as ZF or MMSE detection, using the fact that the error vector between the true transmitted signal vector and its estimate is sparse if the estimate is reliable enough. By using a tentative estimate with some conventional detection, PDSR transforms the original linear equation of the received signal model with the transmitted signal vector into a linear equation with the error vector. If the error vector is sparse, errors can be reconstructed with compressed sensing techniques [7], [8]. In [6], multipath matching pursuit (MMP) [9], which is one of greedy algorithms and is an extension of orthogonal matching pursuit (OMP), is used to estimate the sparse error vector.

This paper proposes a novel error recovery method for massive MIMO signal detection on the basis of the preliminary conference paper [10]. While the conventional PDSR uses the sparsity of the error vector, the proposed method uses the fact that the error is not only sparse but also discrete-valued in practical digital communications systems. To take advantage of the both properties, we firstly formulate the error recovery problem as the maximum a posteriori (MAP) estimation. For large-scale systems, however, the MAP estimation requires a prohibitive computational complexity because it is a combinatorial optimization problem. We thus relax it into the sum-of-absolute-value (SOAV) optimization problem [11], [12] with a similar but slightly different manner compared to [13], which considers the multiuser detection in machine-to-machine communications. While the relaxation in [13] might result in a non-convex optimization problem in general, the proposed relaxation in this paper can always give a convex one. The convex SOAV optimization problem can be efficiently solved with proximal splitting methods [14], such as Beck-Teboulle proximal gradient algorithm and Douglas-Rachford algorithm. To obtain further better performance, we also propose an iterative error recovery, where the estimate obtained in the previous iteration is used as the tentative estimate in each iteration. The proposed method can be applied for binary phase shift keying (BPSK), quadrature phase shift keying (QPSK), and any rectangular quadratic amplitude modulation (QAM). As a theoretical analysis, by using the restricted isometry property (RIP) [15], we give a theoretical upper bound for the size of the reconstruction error, which is defined as the difference between the solution of the SOAV optimization and the true error vector. Simulation results show that the proposed method has better performance than that of the conventional error recovery method for especially large MIMO systems.

In the rest of the paper, we use the following notations:

Manuscript received January 21, 2017.
Manuscript revised June 26, 2017.
†The author is with Graduate School of Informatics, Kyoto University, Kyoto-shi, 606-8501 Japan.
††The author is with Graduate School of Engineering, Osaka City University, Osaka-shi, 558-8585 Japan.
*This paper has been partially presented at ISITA2016, California, USA, October 30–November 2, 2016.
a) E-mail: rhayakawa@sys.i.kyoto-u.ac.jp
b) E-mail: kazunori@eng.osaka-cu.ac.jp
DOI: 10.1587/transfun.E100.A.2671

We denote the set of all real numbers by $\mathbb{R}$ and the set of all complex numbers by $\mathbb{C}$. $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ indicate the real part and the imaginary part, respectively. Superscript $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denote the transpose and the Hermitian transpose, respectively. We represent the imaginary unit by j, the identity matrix by $\boldsymbol{I}$, the $M \times N$ matrix whose elements are all 1 by $\mathbf{1}_{M\times N}$, and the $M \times N$ matrix whose elements are all 0 by $\mathbf{0}_{M\times N}$. For a vector $\boldsymbol{a} = [a_1 \cdots a_N]^{\mathrm{T}} \in \mathbb{R}^N$, the ℓ_0 norm $\|\boldsymbol{a}\|_0$ of $\boldsymbol{a}$ denotes the number of nonzero elements in $\boldsymbol{a}$. We also define the ℓ_1, ℓ_2, and ℓ_∞ norms of $\boldsymbol{a}$ as $\|\boldsymbol{a}\|_1 = \sum_{i=1}^{N} |a_i|$, $\|\boldsymbol{a}\|_2 = \sqrt{\sum_{i=1}^{N} a_i^2}$, and $\|\boldsymbol{a}\|_\infty = \max_{i\in\{1,\dots,N\}} |a_i|$, respectively. For an index set $\mathcal{I} \subset \{1,\dots,N\}$, $\boldsymbol{a}_{\mathcal{I}} \in \mathbb{R}^N$ is defined by

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

where $[\boldsymbol{a}_{\mathcal{I}}]_i$ denotes the ith element of $\boldsymbol{a}_{\mathcal{I}}$. $|\mathcal{I}|$ and $\mathcal{I}^{\mathrm{c}} = \{1,\dots,N\} \setminus \mathcal{I}$ represent the cardinality of $\mathcal{I}$ and the complement set of $\mathcal{I}$, respectively. We denote the Euclidean inner product by $\langle\cdot,\cdot\rangle$.

2. System Model

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

$$\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{s}} + \tilde{\boldsymbol{v}}, \tag{2}$$

where

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

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

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{v}, \tag{4}$$

where

$$\boldsymbol{y} = \begin{bmatrix} \mathrm{Re}\{\tilde{\boldsymbol{y}}\} \\ \mathrm{Im}\{\tilde{\boldsymbol{y}}\} \end{bmatrix} \in \mathbb{R}^{2m},\ \boldsymbol{H} = \begin{bmatrix} \mathrm{Re}\{\tilde{\boldsymbol{H}}\} & -\mathrm{Im}\{\tilde{\boldsymbol{H}}\} \\ \mathrm{Im}\{\tilde{\boldsymbol{H}}\} & \mathrm{Re}\{\tilde{\boldsymbol{H}}\} \end{bmatrix} \in \mathbb{R}^{2m\times 2n},$$
$$\boldsymbol{s} = \begin{bmatrix} \mathrm{Re}\{\tilde{\boldsymbol{s}}\} \\ \mathrm{Im}\{\tilde{\boldsymbol{s}}\} \end{bmatrix} \in \mathbb{S}^{2n},\ \boldsymbol{v} = \begin{bmatrix} \mathrm{Re}\{\tilde{\boldsymbol{v}}\} \\ \mathrm{Im}\{\tilde{\boldsymbol{v}}\} \end{bmatrix} \in \mathbb{R}^{2m}, \tag{5}$$

and $\mathbb{S} = \left\{\mathrm{Re}\{\tilde{x}\} \mid \tilde{x} \in \tilde{\mathbb{S}}\right\} \cup \left\{\mathrm{Im}\{\tilde{x}\} \mid \tilde{x} \in \tilde{\mathbb{S}}\right\}$.

3. Conventional Sparse Error Recovery Method

In the conventional sparse error recovery method [6], the non-sparse system model (2), where $\tilde{\boldsymbol{s}}$ is a dense vector, is converted into the sparse one to apply the compressed sensing technique. Let $\tilde{\boldsymbol{s}}_{\mathrm{est}} \in \mathbb{C}^n$ be a tentative estimate of $\tilde{\boldsymbol{s}}$ obtained by some detection method, e.g., ZF or MMSE detection. When we use the linear MMSE detection, for example, the tentative estimate is given by $\tilde{\boldsymbol{s}}_{\mathrm{est}} = \left(\tilde{\boldsymbol{H}}^{\mathrm{H}}\tilde{\boldsymbol{H}} + \sigma_{\tilde{v}}^2\boldsymbol{I}\right)^{-1}\tilde{\boldsymbol{H}}^{\mathrm{H}}\tilde{\boldsymbol{y}}$. We then obtain $\tilde{\boldsymbol{s}}_{\mathrm{est}}^{\mathrm{d}} = Q_{\tilde{\mathbb{S}}}(\tilde{\boldsymbol{s}}_{\mathrm{est}}) \in \tilde{\mathbb{S}}^n$, where the element-wise function $Q_{\tilde{\mathbb{S}}}(\cdot)$ maps each element into its closest symbol in $\tilde{\mathbb{S}}$, i.e., it provides the hard decision of $\tilde{\boldsymbol{s}}_{\mathrm{est}}$. The key point is that the error vector $\tilde{\boldsymbol{e}} = \tilde{\boldsymbol{s}} - \tilde{\boldsymbol{s}}_{\mathrm{est}}^{\mathrm{d}}$ is sparse if the tentative estimate is reliable enough. The transformation into the sparse system can be performed by subtracting $\tilde{\boldsymbol{H}}\tilde{\boldsymbol{s}}_{\mathrm{est}}^{\mathrm{d}}$ from both sides of (2) as

$$\tilde{\boldsymbol{y}}' \overset{\mathrm{def}}{=} \tilde{\boldsymbol{y}} - \tilde{\boldsymbol{H}}\tilde{\boldsymbol{s}}_{\mathrm{est}}^{\mathrm{d}} = \tilde{\boldsymbol{H}}\left(\tilde{\boldsymbol{s}} - \tilde{\boldsymbol{s}}_{\mathrm{est}}^{\mathrm{d}}\right) + \tilde{\boldsymbol{v}} \tag{6}$$
$$= \tilde{\boldsymbol{H}}\tilde{\boldsymbol{e}} + \tilde{\boldsymbol{v}}. \tag{7}$$

From (7), we can reconstruct the error vector $\tilde{\boldsymbol{e}}$ via some compressed sensing algorithm, such as OMP or MMP. Denoting the estimate of the error vector $\tilde{\boldsymbol{e}}$ as $\tilde{\boldsymbol{e}}_{\mathrm{est}}$, we can obtain the improved estimate of $\tilde{\boldsymbol{s}}$ as $\tilde{\boldsymbol{s}}_{\mathrm{est}}^{\mathrm{d}} + \tilde{\boldsymbol{e}}_{\mathrm{est}}$.

4. Proposed Error Recovery Method

Using the transformation from the non-sparse system into the sparse one, we can reconstruct the error vector via compressed sensing technique. However, the conventional method cannot use the discreteness of the error vector though it is actually not only sparse but also discrete-valued in practical digital communications systems. Moreover, the hard decision $\tilde{\boldsymbol{s}}_{\mathrm{est}}^{\mathrm{d}}$ of the transmitted signal vector instead of the soft decision $\tilde{\boldsymbol{s}}_{\mathrm{est}}$ is used to calculate the error vector, which may result in performance degradation. To achieve better performance, we here propose an error recovery method taking advantage of both sparsity and discreteness as well as the soft decision of the transmitted signal vector. Since the prior distribution of the error is not uniform in general, we firstly consider the MAP estimation of the error vector.

4.1 MAP Estimation

The proposed method uses the real signal model (4) during error recovery. Let $\boldsymbol{s}_{\mathrm{est}} = [s_{\mathrm{est},1} \cdots s_{\mathrm{est},2n}]^{\mathrm{T}} \in \mathbb{R}^{2n}$ be a tentative estimate of $\boldsymbol{s} = [s_1 \cdots s_{2n}]^{\mathrm{T}}$ and $\boldsymbol{s}_{\mathrm{est}}^{\mathrm{d}} = [s_{\mathrm{est},1}^{\mathrm{d}} \cdots s_{\mathrm{est},2n}^{\mathrm{d}}]^{\mathrm{T}} = Q_{\mathbb{S}}(\boldsymbol{s}_{\mathrm{est}}) \in \mathbb{S}^{2n}$ be its hard decision. We firstly transform (4) into

$$\boldsymbol{y}' = \boldsymbol{H}\boldsymbol{e} + \boldsymbol{v}, \tag{8}$$

where $\boldsymbol{y}' = \boldsymbol{y} - \boldsymbol{H}\boldsymbol{s}_{\mathrm{est}}^{\mathrm{d}}$ and $\boldsymbol{e} = \boldsymbol{s} - \boldsymbol{s}_{\mathrm{est}}^{\mathrm{d}}$.

The MAP estimation problem of maximizing $p(\boldsymbol{e} \mid$

$\boldsymbol{y}') \propto p(\boldsymbol{y}' \mid \boldsymbol{e})p(\boldsymbol{e})$ is equivalent to minimizing $-\log p(\boldsymbol{y}' \mid \boldsymbol{e}) - \log p(\boldsymbol{e})$. Since $\boldsymbol{y}'$ is written as (8) and $\boldsymbol{v}$ is the Gaussian noise vector having the covariance matrix of $(\sigma_{\tilde{v}}^2/2)\boldsymbol{I}$, the log likelihood function is given by

$$\log p(\boldsymbol{y}' \mid \boldsymbol{e}) = -\frac{1}{\sigma_{\tilde{v}}^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{e}\|_2^2 - \frac{1}{2}\log(\pi\sigma_{\tilde{v}}^2). \tag{9}$$

Assuming the independence of the elements of $\boldsymbol{e}$, we approximate $p(\boldsymbol{e}) \approx \prod_{j=1}^{2n} p(e_j)$, where e_j represents the jth element of $\boldsymbol{e}$. Thus, the objective function $-\log p(\boldsymbol{y}' \mid \boldsymbol{e}) - \log p(\boldsymbol{e})$ to be minimized for the MAP estimation can be reduced to

$$\frac{1}{\sigma_{\tilde{v}}^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{e}\|_2^2 - \sum_{j=1}^{2n}\log p(e_j). \tag{10}$$

To minimize (10), an explicit expression of $p(e_j)$ will be required. For simplicity, we assume QPSK with $\tilde{\mathbb{S}} = \{1 + \mathrm{j}, -1 + \mathrm{j}, -1 - \mathrm{j}, 1 - \mathrm{j}\}$, $\mathrm{E}[\tilde{\boldsymbol{s}}] = \boldsymbol{0}_{n\times 1}$, and $\mathrm{E}[\tilde{\boldsymbol{s}}\tilde{\boldsymbol{s}}^{\mathrm{H}}] = 2\boldsymbol{I}$. Since $e_j = s_j - s_{\mathrm{est},j}^{\mathrm{d}}$ and $s_j, s_{\mathrm{est},j}^{\mathrm{d}} \in \{1, -1\}$, e_j is discrete-valued and takes a value only in $\mathbb{B} = \{b_0, b_1, b_2\}$, where $b_0 = 0, b_1 = -2, b_2 = 2$. Thus, with the probability $p_{\ell,j} = p(e_j = b_\ell)$ $(\ell = 0, 1, 2)$, $p(e_j)$ can be written as

$$p(e_j) = \prod_{\ell=0}^{2} p_{\ell,j}^{\delta(e_j, b_\ell)} \tag{11}$$

for $e_j = b_\ell$ $(\ell = 0, 1, 2)$, where $\delta(\alpha, \beta) = 1$ if $\alpha = \beta$ and $\delta(\alpha, \beta) = 0$ otherwise, and we define $0^0 = 1$. By substituting (11) into (10), the objective function is rewritten as

$$\frac{1}{\sigma_{\tilde{v}}^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{e}\|_2^2 - \sum_{j=1}^{2n}\sum_{\ell=0}^{2}\delta(e_j, b_\ell)\log p_{\ell,j} \tag{12}$$

$$= \frac{1}{\sigma_{\tilde{v}}^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{e}\|_2^2 - \sum_{j=1}^{2n}\sum_{\ell=0}^{2}(1 - \|e_j - b_\ell\|_0)\log p_{\ell,j}. \tag{13}$$

Hence, the MAP estimation problem can be written as

$$\underset{\boldsymbol{x}\in\mathbb{B}^{2n}}{\text{minimize}}\ \frac{1}{\sigma_{\tilde{v}}^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \sum_{j=1}^{2n}\sum_{\ell=0}^{2}(\log p_{\ell,j})\|x_j - b_\ell\|_0. \tag{14}$$

The probability $p_{\ell,j}$ is given by

$$p_{0,j} = p(s_j = +1),\ p_{1,j} = p(s_j = -1),\ p_{2,j} = 0 \tag{15}$$

if $s_{\mathrm{est},j}^{\mathrm{d}} = +1$, and

$$p_{0,j} = p(s_j = -1),\ p_{1,j} = 0,\ p_{2,j} = p(s_j = +1) \tag{16}$$

if $s_{\mathrm{est},j}^{\mathrm{d}} = -1$. Although both of the prior probabilities $p(s_j = +1)$ and $p(s_j = -1)$ are usually set to be 1/2 when we have no prior information, we can use the tentative estimate $\boldsymbol{s}_{\mathrm{est}}$ as the prior information in the error recovery problem. Hence, we calculate the posterior probabilities $p(s_j = +1 \mid \boldsymbol{y})$ and $p(s_j = -1 \mid \boldsymbol{y})$ by using $\boldsymbol{s}_{\mathrm{est}}$, and substitute them as the prior probabilities in (15) and (16). To obtain the posterior probability, we calculate the posterior log likelihood ratio (LLR) of the transmitted symbols

$$\lambda_j = \log\frac{p(s_j = +1 \mid \boldsymbol{y})}{p(s_j = -1 \mid \boldsymbol{y})} \tag{17}$$

from the estimate $\boldsymbol{s}_{\mathrm{est}}$, which is the soft decision of $\boldsymbol{s}$. For the reduction of the computational complexity, we assume the independence of each received signal and approximate (17) as

$$\lambda_j \approx \sum_{i=1}^{2m}\log\frac{p(y_i \mid s_j = +1)}{p(y_i \mid s_j = -1)}. \tag{18}$$

We further rewrite y_i as $y_i = h_{i,j}s_j + \xi_i^j$, where $\xi_i^j = \sum_{k=1,k\neq j}^{2n} h_{i,k}s_k + v_i$. Regarding ξ_i^j as the Gaussian random variable with mean $\mu_{\xi_i^j}$ and variance $\sigma_{\xi_i^j}^2$ by using the Gaussian approximation [4], we further approximate (18) as

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

$$\approx \sum_{i=1}^{2m}\left[\log\left\{\frac{1}{\sqrt{2\pi\sigma_{\xi_i^j}^2}}\exp\left(-\frac{\left(y_i - h_{i,j} - \mu_{\xi_i^j}\right)^2}{2\sigma_{\xi_i^j}^2}\right)\right\} - \log\left\{\frac{1}{\sqrt{2\pi\sigma_{\xi_i^j}^2}}\exp\left(-\frac{\left(y_i + h_{i,j} - \mu_{\xi_i^j}\right)^2}{2\sigma_{\xi_i^j}^2}\right)\right\}\right] \tag{19}$$

$$= \sum_{i=1}^{2m}\left\{-\frac{\left(y_i - h_{i,j} - \mu_{\xi_i^j}\right)^2}{2\sigma_{\xi_i^j}^2} + \frac{\left(y_i + h_{i,j} - \mu_{\xi_i^j}\right)^2}{2\sigma_{\xi_i^j}^2}\right\} \tag{20}$$

$$= \sum_{i=1}^{2m}\frac{2h_{i,j}\left(y_i - \mu_{\xi_i^j}\right)}{\sigma_{\xi_i^j}^2}, \tag{21}$$

where

$$\mu_{\xi_i^j} = \sum_{k=1,k\neq j}^{2n} h_{i,k}\mathrm{E}[s_k], \tag{22}$$

$$\sigma_{\xi_i^j}^2 = \sum_{k=1,k\neq j}^{2n} h_{i,k}^2\left(1 - \mathrm{E}[s_k]^2\right) + \frac{\sigma_{\tilde{v}}^2}{2}. \tag{23}$$

The expectation $\mathrm{E}[s_k]$ is not available in general. We thus obtain the approximations of $\mu_{\xi_i^j}$ and $\sigma_{\xi_i^j}^2$ by replacing $\mathrm{E}[s_k]$ with

$$s'_{\mathrm{est},k} = \begin{cases} -1 & (s_{\mathrm{est},k} < -1) \\ s_{\mathrm{est},k} & (-1 \le s_{\mathrm{est},k} < 1) \\ 1 & (1 \le s_{\mathrm{est},k}) \end{cases}, \tag{24}$$

which is bounded in $[-1, 1]$, because $s_k \in \{1, -1\}$ and hence $-1 \le \mathrm{E}[s_k] \le 1$. From the approximated posterior LLR $\hat{\lambda}_j$ obtained by using (24), the approximations of the posterior probabilities are given by

$$p(s_j = +1 \mid \boldsymbol{y}) = \frac{e^{\hat{\lambda}_j}}{1 + e^{\hat{\lambda}_j}} = \frac{1}{2}\left\{1 + \tanh\left(\frac{\hat{\lambda}_j}{2}\right)\right\} \tag{25}$$

$$p(s_j = -1 \mid \boldsymbol{y}) = \frac{1}{1 + e^{\hat{\lambda}_j}} = \frac{1}{2}\left\{1 - \tanh\left(\frac{\hat{\lambda}_j}{2}\right)\right\}. \tag{26}$$

4.2 Relaxation into Convex Optimization Problem

Since the problem (14) is the combinatorial optimization problem, it requires a prohibitive computational complexity for large n. We thus consider to relax (14) into a convex optimization problem in the similar way as [13] with the idea used in compressed sensing. However, simple replacements of $\mathbb{B}^{2n}$ and ℓ_0 norm with $\mathbb{R}^{2n}$ and ℓ_1 norm respectively will not necessarily result in the convex problem because $\log p_{\ell,j}$ could be zero or negative. Hence, we firstly replace $\mathbb{B}^{2n}$ and $\log p_{\ell,j}$ with $\mathbb{R}^{2n}$ and $q_{\ell,j} \ge 0$ respectively as

$$\underset{\boldsymbol{x} \in \mathbb{R}^{2n}}{\text{minimize}} \ \frac{1}{\sigma_{\tilde{v}}^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \sum_{j=1}^{2n}\sum_{\ell=0}^{2} q_{\ell,j}\|x_j - b_\ell\|_0, \tag{27}$$

and then relax (27) into the SOAV optimization problem [11], [12] as

$$\underset{\boldsymbol{x} \in \mathbb{R}^{2n}}{\text{minimize}} \ \frac{1}{\sigma_{\tilde{v}}^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \sum_{j=1}^{2n}\sum_{\ell=0}^{2} q_{\ell,j}|x_j - b_\ell|. \tag{28}$$

The coefficients $q_{\ell,j}$ are determined so that they satisfy

$$\sum_{\ell \in \mathcal{L}_j} (\log p_{\ell,j})\|x_j - b_\ell\|_0 + C_j = \sum_{\ell \in \mathcal{L}_j} q_{\ell,j}\|x_j - b_\ell\|_0 \tag{29}$$

for all $x_j = b_\ell$ $(\ell \in \mathcal{L}_j)$, where $\mathcal{L}_j = \{\ell \mid p_{\ell,j} > 0\}$ and C_j is a positive constant. Note that the indices ℓ corresponding to $p_{\ell,j} = 0$ is not considered in the condition (29). For $\ell \notin \mathcal{L}_j$, $q_{\ell,j}$ is fixed to 0. The condition (29) requires that the objective functions in (14) and (27) have the same value for $x_j = b_\ell$ $(\ell \in \mathcal{L}_j)$ up to a constant. For example, if $s^{\mathrm{d}}_{\mathrm{est},j} = 1$, $p_{0,j}, p_{1,j} > 0$ and $p_{2,j} = 0$, then $\mathcal{L}_j = \{0, 1\}$ and the condition (29) becomes $q_{1,j} = \log p_{1,j} + C_j$ and $q_{0,j} = \log p_{0,j} + C_j$. We thus select as $C_j = -\min(\log p_{0,j}, \log p_{1,j}) + \tilde{C}_j$ $(\tilde{C}_j \ge 0)$ and obtain $q_{0,j}, q_{1,j} \ge 0$. In this case, $q_{2,j}$ is fixed to 0. In general, the condition (29) can be written as

$$q_{\ell,j} = \log p_{\ell,j} + \frac{C_j}{|\mathcal{L}_j| - 1} \tag{30}$$

for all $\ell \in \mathcal{L}_j$ (See Appendix A). Hence, we can obtain a nonnegative $q_{\ell,j}$ by selecting C_j as $C_j = -(|\mathcal{L}_j| - 1)\min_{\ell \in \mathcal{L}_j} \log p_{\ell,j} + \tilde{C}_j$ $(\tilde{C}_j \ge 0)$. It should be noted that in the conventional relaxation method [13], the ℓ_0 norm in the right hand of (29) is replaced with the ℓ_1 norm to keep the value of the objective function in (14) and (28) on $\mathbb{B}^{2n}$, except for a constant term. In some cases, however, the optimization problem with the conventional relaxation is still non-convex due to a negative value of $q_{\ell,j}$. On the other hand, the proposed relaxation can always select a positive $q_{\ell,j}$ and ensure that the optimization problem (28) is convex.

The optimization problem (28) can be solved with proximal splitting methods [14]. The improved estimate of the transmitted signal vector $\boldsymbol{s}$ is obtained as $\boldsymbol{s}^{\mathrm{d}}_{\mathrm{est}} + \boldsymbol{e}_{\mathrm{est}}$, where $\boldsymbol{e}_{\mathrm{est}}$ is the minimizer of the problem (28).

4.3 Iterative Error Recovery

To further improve the performance, we also propose an iterative error recovery. In each iteration, the estimate obtained in the previous iteration is used as the tentative estimate. The algorithm of the proposed method is summarized as follows:

Algorithm 1: *(Iterative error recovery via SOAV optimization)*

1. Obtain an initial tentative estimate $\boldsymbol{s}_{\mathrm{est}} \in \mathbb{R}^{2n}$.
2. Iterate a)–e) for T times.
   a. Calculate the approximation of posterior LLR $\hat{\lambda}_j$ from the tentative estimate $\boldsymbol{s}_{\mathrm{est}}$.
   b. Compute $p_{\ell,j}$ with (15), (16), (25), and (26).
   c. Obtain $q_{\ell,j}$ satisfying (29).
   d. Solve the optimization problem (28) and obtain the solution $\boldsymbol{e}_{\mathrm{est}} \in \mathbb{R}^{2n}$.
   e. Modify the tentative estimate into $\boldsymbol{s}^{\mathrm{d}}_{\mathrm{est}} + \boldsymbol{e}_{\mathrm{est}} \in \mathbb{R}^{2n}$.
3. Obtain the final hard decision by applying $Q_{\mathbb{S}}(\cdot)$ to the tentative estimate.

4.4 Extension to other Modulation Schemes

Although we have assumed QPSK modulation so far, we can apply the proposed method for any rectangular QAM symbols. For example, when we use 16-QAM symbols whose real and imaginary parts take +3, +1, −1, or −3, the error e_j takes a value only in $\mathbb{B} = \{0, \pm 2, \pm 4, \pm 6\}$. The corresponding SOAV optimization problem can be obtained by replacing $\sum_{j=1}^{2n}\sum_{\ell=0}^{2} q_{\ell,j}|x_j - b_\ell|$ in (28) with $\sum_{j=1}^{2n}\sum_{\ell=0}^{6} q_{\ell,j}|x_j - b_\ell|$, where $b_0 = 0, b_1 = -6, b_2 = -4, b_3 = -2, b_4 = 2, b_5 = 4$, and $b_6 = 6$. The coefficients $q_{\ell,j}$ can be obtained by using the LLR calculation for 16-QAM symbols [17] and the optimization problem can also be solved via proximal splitting methods. However, for some other modulation methods such as 8-phase shift keying (PSK), we cannot directly apply the proposed method. When we use 8-PSK symbols with the alphabet $\tilde{\mathbb{S}} = \left\{1, \frac{1}{\sqrt{2}} + \mathrm{j}\frac{1}{\sqrt{2}}, \mathrm{j}, -\frac{1}{\sqrt{2}} + \mathrm{j}\frac{1}{\sqrt{2}}, -1, -\frac{1}{\sqrt{2}} - \mathrm{j}\frac{1}{\sqrt{2}}, -\mathrm{j}, \frac{1}{\sqrt{2}} - \mathrm{j}\frac{1}{\sqrt{2}}\right\}$, the proposed method may provide inappropriate estimates, such as $1 - \mathrm{j}\frac{1}{\sqrt{2}}$, 0, and $\frac{1}{\sqrt{2}}$.

5. Performance Analysis

In this section, we give a theoretical performance analysis for the reconstruction of a discrete-valued sparse vector via the SOAV optimization. We use RIP considered in the performance analysis for the reconstruction of a sparse vector via compressed sensing [15].

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

Definition 2 (RIP): A matrix $\boldsymbol{\Phi}$ satisfies RIP of order K if there is a constant $\delta_K \in (0, 1)$ such that

$$(1 - \delta_K)\|\boldsymbol{x}\|_2^2 \leq \|\boldsymbol{\Phi x}\|_2^2 \leq (1 + \delta_K)\|\boldsymbol{x}\|_2^2 \tag{31}$$

holds for all K-sparse vector $\boldsymbol{x}$. The minimum value of possible δ_K is called K-restricted isometry constant.

In this section, we consider the reconstruction of a discrete-valued sparse vector $\boldsymbol{e} \in \mathbb{B}^N = \{b_0, b_1, \ldots, b_L\}^N$ ($b_0 = 0$) from its linear measurements $\boldsymbol{y}' = \boldsymbol{He} + \boldsymbol{v}$. The SOAV optimization problem for the reconstruction is given by

$$\underset{\boldsymbol{x} \in \mathbb{R}^N}{\text{minimize}} \ \frac{1}{\sigma_{\tilde{v}}^2}\|\boldsymbol{y}' - \boldsymbol{Hx}\|_2^2 + \sum_{j=1}^{N}\sum_{\ell=0}^{L} q_{\ell,j}|x_j - b_\ell|, \tag{32}$$

which is a generalization of (28). The optimization problem (32) is equivalent to

$$\begin{aligned} &\underset{\boldsymbol{x} \in \mathbb{R}^N}{\text{minimize}} \ \sum_{j=1}^{N}\sum_{\ell=0}^{L} q_{\ell,j}|x_j - b_\ell| \\ &\text{subject to } \|\boldsymbol{y}' - \boldsymbol{Hx}\|_2 \leq \varepsilon \end{aligned} \tag{33}$$

with a proper choice of the constant $\varepsilon > 0$ corresponding to $\sigma_{\tilde{v}}^2$. For the solution of (33), we have the following theorem:

Theorem 1: Let $\boldsymbol{e}_{\text{est}}$ be the solution of (33). Assume that the true vector $\boldsymbol{e}$ is K-sparse and satisfies the constraint in (33), i.e., $\|\boldsymbol{y}' - \boldsymbol{He}\|_2 \leq \varepsilon$. We define δ_{2K} as the $2K$-restricted isometry constant of $\boldsymbol{H}$, $\mathcal{J}_\ell = \{j \mid e_j = b_\ell\}$, $Q_j = \sum_{\ell=0}^{L} q_{\ell,j}$ and $Q_{\min} = \min_{j \in \mathcal{J}_0} Q_j$. If the inequality

$$\delta_{2K} < \frac{1}{\sqrt{2}c + 1} \tag{34}$$

holds, then we have

$$\|\boldsymbol{e}_{\text{est}} - \boldsymbol{e}\|_2 \leq \frac{1}{1 - c\rho}\left\{(1 + c)\tau\varepsilon + \frac{2}{\sqrt{K}}(1 + \rho)I\right\}, \tag{35}$$

where

$$\begin{aligned} \tau &= \frac{2\sqrt{1 + \delta_{2K}}}{1 - \delta_{2K}},\ \rho = \frac{\sqrt{2}\delta_{2K}}{1 - \delta_{2K}},\ I = \sum_{j \in \mathcal{J}_0}\sum_{\ell=1}^{L}\frac{q_{\ell,j}}{Q_{\min}}|b_\ell|, \\ c &= \sqrt{\frac{\sum_{\ell=1}^{L}\sum_{j \in \mathcal{J}_\ell} r_{\ell,j}^2}{K}},\ r_{\ell,j} = \max_\ell\left(\frac{Q_j - q_{\ell,j}}{Q_{\min}}, 0\right). \end{aligned} \tag{36}$$

Proof. See Appendix B. □

Theorem 1 can be considered as a generalization of the performance analysis for the reconstruction of K-sparse vector via ℓ_1 optimization [15]. Actually, if $q_{0,j} = 1$ and $q_{\ell,j} = 0$ ($\ell = 1, \ldots, L$) for all j, then (33) is written as the ℓ_1 optimization problem

$$\underset{\boldsymbol{x} \in \mathbb{R}^N}{\text{minimize}} \ \sum_{j=1}^{N}|x_j| \text{ subject to } \|\boldsymbol{y}' - \boldsymbol{Hx}\|_2 \leq \varepsilon. \tag{37}$$

Since $c = 1$ and $I = 0$ in this case, the condition (34) and the upper bound (35) can be written as

$$\delta_{2K} < \frac{1}{\sqrt{2} + 1} = \sqrt{2} - 1, \tag{38}$$

$$\|\boldsymbol{e}_{\text{est}} - \boldsymbol{e}\|_2 \leq \frac{2}{1 - \rho}\tau\varepsilon, \tag{39}$$

respectively, which correspond to the result for ℓ_1 optimization.

The condition (34) can be milder than (38) for an appropriate choice of $q_{\ell,j}$. Since $\boldsymbol{e}$ is K-sparse, we have $\sum_{\ell=1}^{L}|\mathcal{J}_\ell| \leq K$. Thus, in the case of $Q_j = 1$ ($j = 1, \ldots, N$), for example, $c \leq 1$ follows from $r_{\ell,j} = \max(1 - q_{\ell,j}, 0) \leq 1$. If $q_{\ell,j} > 0$ for some $\ell \geq 1$ and $j \in \mathcal{J}_\ell$, then we have $r_{\ell,j} < 1$ and $c < 1$. In this case, the condition (34) is milder than (38) because $1/(\sqrt{2}c + 1) > \sqrt{2} - 1$. Also the upper bound (35) for $c < 1$ is smaller than (39) because it is the monotonically increasing function of c.

Unfortunately, it is difficult in general to obtain the restricted isometry constant δ_K for a specific matrix because of the infeasible computational complexity. For random matrices, however, some asymptotical results about RIP have been obtained [16] and hence they might be used with Theorem 1 for very large MIMO systems. Moreover, Theorem 1 may provide different criterion for calculating the coefficients $q_{\ell,j}$ from the method described in Section 4.

6. Simulation Results

In this section, we evaluate the bit error rate (BER) performance of the proposed method and the conventional method [6] via computer simulations. In the simulations, flat Rayleigh fading channels are assumed and $\tilde{\boldsymbol{H}}$ is composed of i.i.d. complex Gaussian random variables with zero mean and unit variance. We transmit 100 symbol vectors for each realization of $\tilde{\boldsymbol{H}}$, and obtain the average BER over the

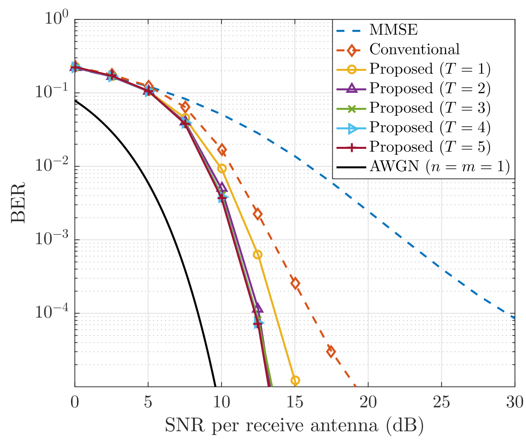

Fig. 1 BER performance for $(n, m) = (32, 32)$.

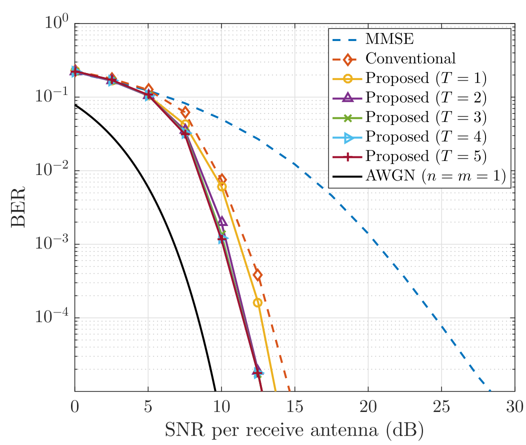

Fig. 2 BER performance for $(n, m) = (128, 128)$.

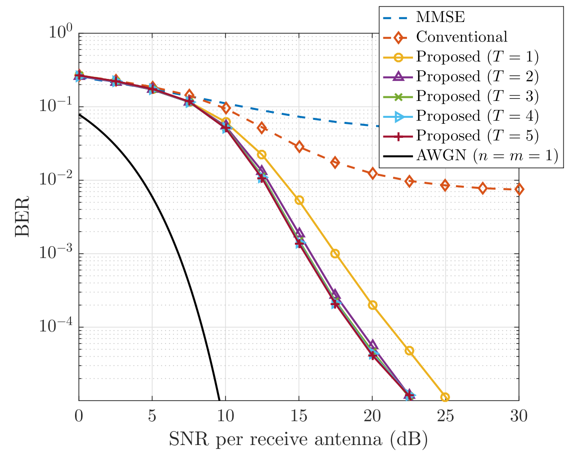

Fig. 3 BER performance for $(n, m) = (32, 24)$.

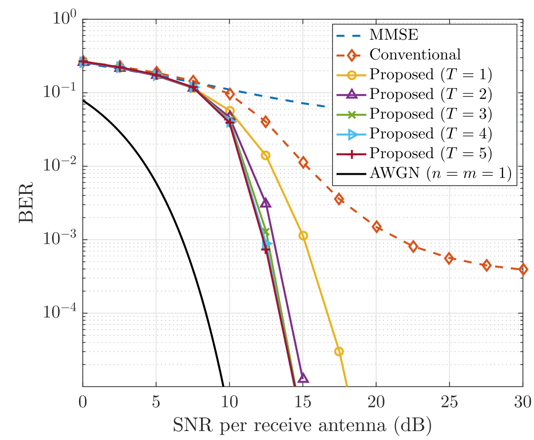

Fig. 4 BER performance for $(n, m) = (128, 96)$.

transmission of 6.4×10^6 symbols. The modulation scheme is QPSK. For the MMP algorithm in the conventional method, the number of iterations, the number of paths from each candidate in each iteration, and the maximum number of candidates in each iteration are set to $K_{\mathrm{MMP}} = \lceil 0.15n \rceil$, $L_{\mathrm{MMP}} = 2$, $N_{\mathrm{MMP}} = 5$, respectively. In the proposed relaxation for the MAP estimation, $\tilde{C}_j = 1$ $(j = 1, \ldots, 2n)$ is used. The parameters of the Douglas-Rachford algorithm [14] to solve (28) are set to be the same as in [10].

Figures 1 and 2 show the BER performance for the MIMO systems with $(n, m) = (32, 32), (128, 128)$, respectively. "MMSE", "Conventional", and "Proposed" denote linear MMSE detection, conventional error recovery method via MMP, and the proposed error recovery method via SOAV optimization, respectively. For comparison, we also show the theoretical BER curve for the additive white Gaussian noise (AWGN) channel as "AWGN $(n = m = 1)$". In both error recovery methods, the estimate of MMSE detection is used as the initial estimate. As described in Algorithm 1, T indicates the number of iterations of the error recovery in the proposed method. The figures show that the proposed method outperforms the conventional method even when $T = 1$. It is because the proposed method uses the discreteness of the error vector and the initial soft decision, which are not considered in the conventional method. We can also see that the performance is further improved by iterating the error recovery. One of major possible reasons for the performance difference between Figs. 1 and 2 is the channel hardening effect [18], which means that the off-diagonal elements of $\tilde{\boldsymbol{H}}^{\mathrm{H}}\tilde{\boldsymbol{H}}$ become negligible compared to the diagonal elements as the number of antennas increases.

Figures 3 and 4 show the performance for $(n, m) = (32, 24), (128, 96)$, respectively. Such scenario, where the number of receive antennas is less than that of transmitted streams, is called overloaded (or underdetermined) MIMO [19]. Since the performance of MMSE detection is severely degraded in overloaded MIMO, the conventional method also has a poor performance. However, the proposed method performs well in this case as well, especially in large-scale systems. The difference of the performance between Figs. 3 and 4 may be partly caused by the accuracy of the Gaussian approximation in the calculation of the posterior LLRs.

To compare the computational complexity, we evaluate the average computation time to detect a transmitted signal vector and the corresponding BER performance versus n $(= m)$ for the SNR per receive antenna of 12.5 dB in

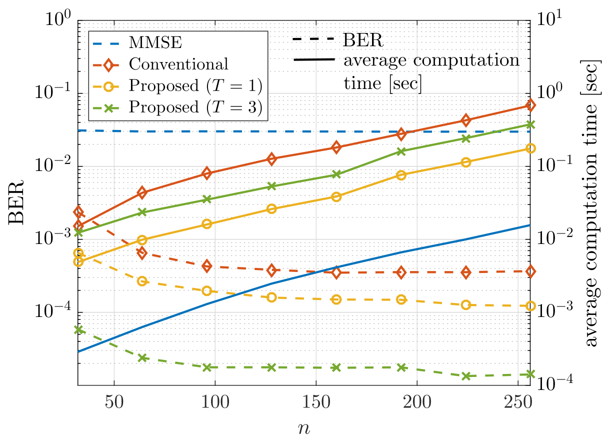

Fig. 5 BER performance and average computation time versus n $(= m)$ for SNR per receive antenna of 12.5 dB.

Fig. 5. The simulation is conducted by using a computer with 2 GHz Intel Core i7-3667U and 8 GB memory. We can see that the proposed error recovery method can achieve better BER performance with lower complexity compared to the conventional method.

7. Conclusion

In this paper, we have proposed the error recovery method for massive MIMO signal detection. The proposed method estimates the error vector via the SOAV optimization, which can take advantage of both the sparsity and the discreteness of the error vector. We have provided the theoretical performance analysis for the reconstruction of the discrete-valued sparse vector via the SOAV optimization. Simulation results show that the proposed method outperforms the conventional error recovery method. Future work includes the extension of the proposed method for other modulation methods as well as the direct reconstruction of complex signals.

Acknowledgments

This work was supported in part by the Grants-in-Aid for Scientific Research no. 15K06064 and 15H2252 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] 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, April 2008.
[3] 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.
[4] T. Wo and P.A. Hoeher, "A simple iterative Gaussian detector for severely delay-spread MIMO channels," Proc. IEEE ICC 2007, pp.4598–4563, June 2007.
[5] R. Hayakawa, K. Hayashi, H. Sasahara, and M. Nagahara, "Massive overloaded MIMO signal detection via convex optimization with proximal splitting," Proc. EUSIPCO 2016, pp.1383–1387, Aug.–Sept. 2016.
[6] J. Choi and B. Shim, "New approach for massive MIMO detection using sparse error recovery," Proc. IEEE GLOBECOM 2014, pp.3754–3759, Dec. 2014.
[7] D.L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol.52, no.4, pp.1289–1306, April 2006.
[8] K. Hayashi, M. Nagahara, and T. Tanaka, "A user's guide to compressed sensing for communications systems," IEICE Trans. Commun., vol.E96-B, no.3, pp.685–712, March 2013.
[9] S. Kwon, J. Wang, and B. Shim, "Multipath matching pursuit," IEEE Trans. Inf. Theory, vol.60, no.5, pp.2986–3001, May 2014.
[10] R. Hayakawa and K. Hayashi, "Error recovery with relaxed MAP estimation for massive MIMO signal detection," Proc. ISITA 2016, pp.483–487, Oct.–Nov. 2016.
[11] M. Nagahara, "Discrete signal reconstruction by sum of absolute values," IEEE Signal Process. Lett., vol.22, no.10, pp.1575–1579, Oct. 2015.
[12] A.A. 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, April 2015.
[13] H. Sasahara, K. Hayashi, and M. Nagahara, "Multiuser detection by MAP estimation with sum-of-absolute-values relaxation," Proc. IEEE ICC 2016, May 2016.
[14] P. Combettes and J. Pesquet, "Proximal splitting methods in signal processing," Fixed-Point Algorithms for Inverse Problems in Science and Engineering, ser. Springer Optimization and Its Applications, vol.49, pp.185–212, Springer New York, 2011.
[15] 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.
[16] 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.
[17] S. Allpress, C. Luschi, and S. Felix, "Exact and approximated expressions of the log-likelihood ratio for 16-QAM signals," Proc. 38th Asilomar Conference on Signal, System and Computer, Nov. 2004.
[18] B.M. Hochwald, T.L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," IEEE Trans. Inf. Theory, vol.50, no.9, pp.1893–1909, Sept. 2004.
[19] 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.

Appendix A: Derivation of (30) from (29)

Let $\mathcal{L}_j = \{\ell_1, \ldots, \ell_{|\mathcal{L}_j|}\}$. For $x_j = b_k$ $(k = \ell_1, \ldots, \ell_{|\mathcal{L}_j|})$, (29) is written as

$$\left(\sum_{\ell \in \mathcal{L}_j \setminus \{k\}} \log p_{\ell,j} \right) + C_j = \sum_{\ell \in \mathcal{L}_j \setminus \{k\}} q_{\ell,j}. \qquad \text{(A· 1)}$$

The conditions (A· 1) for all $k = \ell_1, \ldots, \ell_{|\mathcal{L}_j|}$ are written as

$$\boldsymbol{\Theta}\boldsymbol{\pi} + C_j \mathbf{1}_{|\mathcal{L}_j| \times 1} = \boldsymbol{\Theta}\boldsymbol{q}, \qquad \text{(A· 2)}$$

where $\boldsymbol{\Theta} = \mathbf{1}_{|\mathcal{L}_j| \times |\mathcal{L}_j|} - \boldsymbol{I}$, $\boldsymbol{\pi} = [\log p_{\ell_1,j} \ \cdots \ \log p_{\ell_{|\mathcal{L}_j|},j}]^{\mathrm{T}}$, and $\boldsymbol{q} = [q_{\ell_1,j} \ \cdots \ q_{\ell_{|\mathcal{L}_j|},j}]^{\mathrm{T}}$. Since $\boldsymbol{\Theta}^{-1} = \mathbf{1}_{|\mathcal{L}_j|\times|\mathcal{L}_j|}/(|\mathcal{L}_j|-1) - \boldsymbol{I}$, we have

$$\boldsymbol{q} = \boldsymbol{\pi} + C_j \boldsymbol{\Theta}^{-1}\mathbf{1}_{|\mathcal{L}_j|\times 1} \tag{A·3}$$

$$= \boldsymbol{\pi} + \frac{C_j}{|\mathcal{L}_j|-1}\mathbf{1}_{|\mathcal{L}_j|\times 1}. \tag{A·4}$$

Appendix B: Proof of Theorem 1

Let $\boldsymbol{\xi} = [\xi_1 \ \cdots \ \xi_N]^{\mathrm{T}} = \boldsymbol{e}_{\mathrm{est}} - \boldsymbol{e}$. We define $\mathcal{T}_1$ as the set of indices corresponding to K largest elements in $\boldsymbol{\xi}_{\mathcal{J}_0}$. Similarly, $\mathcal{T}_2$ indicates the set of indices corresponding to K largest elements in $\boldsymbol{\xi}_{(\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1)^{\mathrm{c}}}$, and $\mathcal{T}_3$ indicates that for $\boldsymbol{\xi}_{(\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1\cup\mathcal{T}_2)^{\mathrm{c}}}$. We also define $\mathcal{T}_4, \mathcal{T}_5, \ldots$ in the same manner. Note that the vectors $\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}}, \boldsymbol{\xi}_{\mathcal{T}_1}, \boldsymbol{\xi}_{\mathcal{T}_2}, \ldots$ are all K-sparse, the sets $\mathcal{J}_0^{\mathrm{c}}, \mathcal{T}_1, \mathcal{T}_2, \ldots$ are disjoint, and $\mathcal{J}_0 = \bigcup_{u\geq 1}\mathcal{T}_u$. $\|\boldsymbol{\xi}\|_2$ is upper bounded as

$$\|\boldsymbol{\xi}\|_2 \leq \|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2 + \sum_{u\geq 2}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2. \tag{A·5}$$

First, we evaluate $\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2$. Since $\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}$ is $2K$-sparse, we obtain

$$\begin{aligned}&(1-\delta_{2K})\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2^2\\ &\leq \|\boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2^2 &\text{(A·6)}\\ &= \left\langle \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}, \boldsymbol{H}\boldsymbol{\xi}\right\rangle - \left\langle \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}, \sum_{u\geq 2}\boldsymbol{H}\boldsymbol{\xi}_{\mathcal{T}_u}\right\rangle &\text{(A·7)}\\ &\leq \left|\left\langle \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}, \boldsymbol{H}\boldsymbol{\xi}\right\rangle\right| + \left|\left\langle \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}, \sum_{u\geq 2}\boldsymbol{H}\boldsymbol{\xi}_{\mathcal{T}_u}\right\rangle\right|. &\text{(A·8)}\end{aligned}$$

The first term in (A·8) is bounded as

$$\left|\left\langle \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}, \boldsymbol{H}\boldsymbol{\xi}\right\rangle\right| \leq \|\boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2\|\boldsymbol{H}\boldsymbol{\xi}\|_2. \tag{A·9}$$

From the inequalities of

$$\|\boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2^2 \leq (1+\delta_{2K})\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2^2 \tag{A·10}$$

and

$$\begin{aligned}\|\boldsymbol{H}\boldsymbol{\xi}\|_2 &= \|\boldsymbol{H}(\boldsymbol{e}_{\mathrm{est}}-\boldsymbol{e})\|_2 &\text{(A·11)}\\ &\leq \|\boldsymbol{H}\boldsymbol{e}_{\mathrm{est}}-\boldsymbol{y}'\|_2 + \|\boldsymbol{H}\boldsymbol{e}-\boldsymbol{y}'\|_2 &\text{(A·12)}\\ &\leq 2\varepsilon, &\text{(A·13)}\end{aligned}$$

we have

$$\left|\left\langle \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}, \boldsymbol{H}\boldsymbol{\xi}\right\rangle\right| \leq 2\varepsilon\sqrt{1+\delta_{2K}}\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2. \tag{A·14}$$

The second term in (A·8) is bounded as

$$\begin{aligned}&\left|\left\langle \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}, \sum_{u\geq 2}\boldsymbol{H}\boldsymbol{\xi}_{\mathcal{T}_u}\right\rangle\right|\\ &\leq \sum_{u\geq 2}\left(\left|\left\langle \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}}, \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{T}_u}\right\rangle\right| + |\langle \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{T}_1}, \boldsymbol{H}\boldsymbol{\xi}_{\mathcal{T}_u}\rangle|\right) &\text{(A·15)}\\ &\leq \sum_{u\geq 2}\left(\delta_{2K}\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}}\|_2\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2 + \delta_{2K}\|\boldsymbol{\xi}_{\mathcal{T}_1}\|_2\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2\right) &\text{(A·16)}\\ &= \delta_{2K}\left(\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}}\|_2 + \|\boldsymbol{\xi}_{\mathcal{T}_1}\|_2\right)\sum_{u\geq 2}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2 &\text{(A·17)}\\ &\leq \sqrt{2}\delta_{2K}\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2\sum_{u\geq 2}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2. &\text{(A·18)}\end{aligned}$$

From (A·8), (A·14), and (A·18), we have

$$\begin{aligned}&(1-\delta_{2K})\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2^2\\ &\leq 2\varepsilon\sqrt{1+\delta_{2K}}\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2\\ &\quad + \sqrt{2}\delta_{2K}\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2\sum_{u\geq 2}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2. &\text{(A·19)}\end{aligned}$$

Dividing both sides with $(1-\delta_{2K})\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2$, we obtain

$$\|\boldsymbol{\xi}_{\mathcal{J}_0^{\mathrm{c}}\cup\mathcal{T}_1}\|_2 \leq \tau\varepsilon + \rho\sum_{u\geq 2}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2, \tag{A·20}$$

where τ and ρ are defined in (36).

Next, we evaluate $\sum_{u\geq 2}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2$, which appears in (A·5) and (A·20). For $u \geq 2$, we have

$$\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2 \leq \sqrt{K}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_\infty \leq \frac{1}{\sqrt{K}}\|\boldsymbol{\xi}_{\mathcal{T}_{u-1}}\|_1, \tag{A·21}$$

and thus

$$\sum_{u\geq 2}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2 \leq \frac{1}{\sqrt{K}}\sum_{u\geq 1}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_1 = \frac{1}{\sqrt{K}}\|\boldsymbol{\xi}_{\mathcal{J}_0}\|_1. \tag{A·22}$$

Then, we evaluate $\|\boldsymbol{\xi}_{\mathcal{J}_0}\|_1$ by using the fact that $\boldsymbol{e}_{\mathrm{est}} = \boldsymbol{e} + \boldsymbol{\xi}$ is the solution of the optimization problem (33). Since $\boldsymbol{e}_{\mathrm{est}} = \boldsymbol{e} + \boldsymbol{\xi}$ is the minimizer of the objective function, we have

$$\begin{aligned}&\sum_{j=1}^{2n}\sum_{\ell=0}^{L} q_{\ell,j}|e_j - b_\ell|\\ &\geq \sum_{j=1}^{2n}\sum_{\ell=0}^{L} q_{\ell,j}|e_j + \xi_j - b_\ell| &\text{(A·23)}\\ &= \sum_{j\in\mathcal{J}_0^{\mathrm{c}}}\sum_{\ell=0}^{L} q_{\ell,j}|e_j + \xi_j - b_\ell| + \sum_{j\in\mathcal{J}_0}\sum_{\ell=0}^{L} q_{\ell,j}|e_j + \xi_j - b_\ell| &\text{(A·24)}\\ &\geq \sum_{j\in\mathcal{J}_0^{\mathrm{c}}}\sum_{\ell=0}^{L} q_{\ell,j}|e_j + \xi_j - b_\ell|\\ &\quad + \sum_{j\in\mathcal{J}_0}\sum_{\ell=0}^{L} q_{\ell,j}(|\xi_j| - |e_j - b_\ell|) &\text{(A·25)}\\ &= \sum_{j\in\mathcal{J}_0^{\mathrm{c}}}\sum_{\ell=0}^{L} q_{\ell,j}|e_j + \xi_j - b_\ell|\\ &\quad + \sum_{j\in\mathcal{J}_0}\left(Q_j|\xi_j| - \sum_{\ell=0}^{L} q_{\ell,j}|e_j - b_\ell|\right), &\text{(A·26)}\end{aligned}$$

which gives

$$\sum_{j\in\mathcal{J}_0} Q_j|\xi_j| \le \sum_{j\in\mathcal{J}_0^{\rm c}}\sum_{\ell=0}^{L} q_{\ell,j}(|e_j-b_\ell|-|e_j+\xi_j-b_\ell|) + 2\sum_{j\in\mathcal{J}_0}\sum_{\ell=0}^{L} q_{\ell,j}|e_j-b_\ell|. \quad (\text{A}\cdot 27)$$

The first term in (A· 27) can be bounded as

$$\sum_{j\in\mathcal{J}_0^{\rm c}}\sum_{\ell=0}^{L} q_{\ell,j}(|e_j-b_\ell|-|e_j+\xi_j-b_\ell|)$$

$$= \sum_{k=1}^{L}\sum_{j\in\mathcal{J}_k}\sum_{\ell=0}^{L} q_{\ell,j}(|e_j-b_\ell|-|e_j+\xi_j-b_\ell|) \quad (\text{A}\cdot 28)$$

$$= \sum_{k=1}^{L}\sum_{j\in\mathcal{J}_k}\sum_{\ell=0}^{L} q_{\ell,j}(|b_k-b_\ell|-|b_k+\xi_j-b_\ell|) \quad (\text{A}\cdot 29)$$

$$= \sum_{k=1}^{L}\sum_{j\in\mathcal{J}_k}\left\{-q_{k,j}|\xi_j| + \sum_{\ell\ne k} q_{\ell,j}(|b_k-b_\ell|-|b_k+\xi_j-b_\ell|)\right\} \quad (\text{A}\cdot 30)$$

$$\le \sum_{k=1}^{L}\sum_{j\in\mathcal{J}_k}\left(-q_{k,j}|\xi_j| + \sum_{\ell\ne k} q_{\ell,j}|\xi_j|\right) \quad (\text{A}\cdot 31)$$

$$= \sum_{k=1}^{L}\sum_{j\in\mathcal{J}_k}\left(Q_j - 2q_{k,j}\right)|\xi_j|. \quad (\text{A}\cdot 32)$$

Since $e_j = 0$ for $j\in\mathcal{J}_0$, the second term in (A· 27) is written as

$$2\sum_{j\in\mathcal{J}_0}\sum_{\ell=0}^{L} q_{\ell,j}|e_j-b_\ell| = 2\sum_{j\in\mathcal{J}_0}\sum_{\ell=0}^{L} q_{\ell,j}|b_\ell|. \quad (\text{A}\cdot 33)$$

From (A· 27), (A· 32), (A· 33), and $Q_{\min} \le Q_j$ $(j\in\mathcal{J}_0)$, we obtain

$$\|\boldsymbol{\xi}_{\mathcal{J}_0}\|_1 \le \frac{1}{Q_{\min}}\sum_{j\in\mathcal{J}_0} Q_j|\xi_j| \quad (\text{A}\cdot 34)$$

$$\le \sum_{k=1}^{L}\sum_{j\in\mathcal{J}_k}\frac{Q_j-2q_{k,j}}{Q_{\min}}|\xi_j| + 2\sum_{j\in\mathcal{J}_0}\sum_{\ell=0}^{L}\frac{q_{\ell,j}}{Q_{\min}}|b_\ell|. \quad (\text{A}\cdot 35)$$

By using the Cauchy-Schwarz inequality, we have

$$\sum_{k=1}^{L}\sum_{j\in\mathcal{J}_k}\frac{Q_j-2q_{k,j}}{Q_{\min}}|\xi_j| \le \sqrt{\sum_{k=1}^{L}\sum_{j\in\mathcal{J}_k} r_{k,j}^2}\,\|\boldsymbol{\xi}_{\mathcal{J}_0^{\rm c}}\|_2 \quad (\text{A}\cdot 36)$$

$$= c\sqrt{K}\|\boldsymbol{\xi}_{\mathcal{J}_0^{\rm c}}\|_2, \quad (\text{A}\cdot 37)$$

where $r_{k,j}$ and c are defined in (36). From (A· 35) and (A· 37), $\|\boldsymbol{\xi}_{\mathcal{J}_0}\|_1$ is bounded as

$$\|\boldsymbol{\xi}_{\mathcal{J}_0}\|_1 \le c\sqrt{K}\|\boldsymbol{\xi}_{\mathcal{J}_0^{\rm c}}\|_2 + 2I, \quad (\text{A}\cdot 38)$$

where I is defined in (36). Substituting (A· 38) into (A· 22) gives

$$\sum_{u\ge 2}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2 \le c\|\boldsymbol{\xi}_{\mathcal{J}_0^{\rm c}}\|_2 + \frac{2}{\sqrt{K}}I \quad (\text{A}\cdot 39)$$

$$\le c\|\boldsymbol{\xi}_{\mathcal{J}_0^{\rm c}\cup\mathcal{T}_1}\|_2 + \frac{2}{\sqrt{K}}I. \quad (\text{A}\cdot 40)$$

From (A· 20) and (A· 40), we have

$$\|\boldsymbol{\xi}_{\mathcal{J}_0^{\rm c}\cup\mathcal{T}_1}\|_2 \le \tau\varepsilon + c\rho\|\boldsymbol{\xi}_{\mathcal{J}_0^{\rm c}\cup\mathcal{T}_1}\|_2 + \frac{2}{\sqrt{K}}\rho I. \quad (\text{A}\cdot 41)$$

If $1 - c\rho > 0$, i.e., $\delta_{2K} < 1/(\sqrt{2}c+1)$, then (A· 41) can be written as

$$\|\boldsymbol{\xi}_{\mathcal{J}_0^{\rm c}\cup\mathcal{T}_1}\|_2 \le \frac{1}{1-c\rho}\left(\tau\varepsilon + \frac{2}{\sqrt{K}}\rho I\right), \quad (\text{A}\cdot 42)$$

and hence we obtain

$$\sum_{u\ge 2}\|\boldsymbol{\xi}_{\mathcal{T}_u}\|_2 \le \frac{c}{1-c\rho}\left(\tau\varepsilon + \frac{2}{\sqrt{K}}\rho I\right) + \frac{2}{\sqrt{K}}I \quad (\text{A}\cdot 43)$$

$$\le \frac{1}{1-c\rho}\left(c\tau\varepsilon + \frac{2}{\sqrt{K}}I\right) \quad (\text{A}\cdot 44)$$

from (A· 40). We conclude from (A· 5), (A· 42), and (A· 44) that

$$\|\boldsymbol{\xi}\|_2 \le \frac{1}{1-c\rho}\left\{(1+c)\tau\varepsilon + \frac{2}{\sqrt{K}}(1+\rho)I\right\}. \quad (\text{A}\cdot 45)$$

Ryo Hayakawa received the B.E. and M.E. degrees 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 IEEE.

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 IEEE, APSIPA, and ISCIE.