Error Recovery with Relaxed MAP Estimation for Massive MIMO Signal Detection

Ryo Hayakawa and Kazunori Hayashi
Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606–8501, Japan
Email: rhayakawa@sys.i.kyoto-u.ac.jp, kazunori@i.kyoto-u.ac.jp

***Abstract*—This paper proposes a maximum a posteriori (MAP) estimation-based error recovery method for massive multiple-input multiple-output (MIMO) signal detection. The error recovery is a technique to improve the estimate of transmitted signals taking advantage of the sparsity of the error signal. We formulate the error recovery problem as the MAP estimation, where not only the sparsity but also the discreteness of the error are taken into consideration explicitly. In the proposed MAP estimation, we can also use not only the hard decision of the transmitted signal vector but also the soft decision obtained before the error recovery. The problem of MAP estimation is relaxed into the sum-of-absolute-value optimization problem, which can be efficiently solved with proximal splitting methods. Simulation results show that the proposed method outperforms the conventional method in terms of bit error rate (BER) performance.**

I. Introduction

To achieve very high spectral efficiency in wireless communication systems, much focus has been placed on massive multiple-input multiple-output (MIMO) communications, where tens or hundreds of antennas are equipped in each transmitter and receiver [1]. Since the computational complexity of the signal detection for MIMO systems generally increases along with the increase of the antennas, a low-complexity detection scheme will be required for such massive MIMO communications. Linear signal detections, including the zero forcing (ZF) and the minimum mean square error (MMSE) detection, can be possible candidates for massive MIMO systems. However, since the performance of the linear detectors is far from the optimal maximum likelihood (ML) detection, some non-linear detection schemes have also been proposed to achieve nearly optimal performance. The likelihood ascent search (LAS) [2] and the reactive tabu search (RTS) [3] are non-linear detectors based on the local neighborhood search of likelihood. The belief propagation techniques also provide a low complexity detector [4].

As another approach for massive MIMO signal detection, post-detection sparse error recovery (PDSR) has been proposed [5]. It improves the estimate obtained by the conventional detections, 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 the initial estimate with some conventional detections, PDSR transforms the original linear equation of the transmitted signal vector into a linear equation of the error vector. If the error vector is sparse, errors can be specified with compressed sensing technique [6]. In [5], multipath matching pursuit (MMP) [7], which is a greedy algorithm and an extension of orthogonal matching pursuit (OMP), is used as the compressed sensing algorithm.

In this paper, we propose an error recovery method for massive MIMO systems, which uses the fact that the error vector is not only sparse but also discrete when we employ digital modulations. To use the both properties of the error vector, we formulate the error recovery as a maximum a posteriori (MAP) estimation. In the formulation, we use the initial soft decision of the transmitted signal vector obtained before the error recovery, while the conventional recovery method uses the hard decision alone. Since the MAP estimation for the error vector is a combinational optimization problem, we relax it into the sum-of-absolute-value (SOAV) optimization problem [8] with a similar approach as in [9], which has been proposed for the multiuser detection in machine-to-machine communications. While the conventional relaxation might result in a non-convex optimization problem in general, the proposed relaxation method can always give a convex one. The convex SOAV optimization problem after the relaxation can be solved with the low-complexity proximal splitting methods [10]. Simulation results show that the proposed method outperforms the conventional method for large and very large MIMO systems.

In the rest of the paper, we use the following notations: 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 $\boldsymbol{I}$, and a vector whose elements are all 0 by $\boldsymbol{0}$. For a vector $\boldsymbol{a} = [a_1 \cdots a_N]^{\mathrm{T}} \in \mathbb{R}^N$, ℓ_0 norm $\|\boldsymbol{a}\|_0$ of $\boldsymbol{a}$ denotes the number of nonzero elements in $\boldsymbol{a}$. We also define the ℓ_1 and ℓ_2 norms of $\boldsymbol{a}$ as $\|\boldsymbol{a}\|_1 = \sum_{i=1}^{N} |a_i|$ and $\|\boldsymbol{a}\|_2 = \sqrt{\sum_{i=1}^{N} a_i^2}$, respectively.

II. 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. In addition, we employ quadrature phase shift keying (QPSK) and define the alphabet of the transmitted symbols as $\tilde{\mathbb{S}} = \{1+\mathrm{j}, -1+\mathrm{j}, -1-\mathrm{j}, 1-\mathrm{j}\}$. 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 $\tilde{s}_j$ $(j = 1, \ldots, n)$ denotes the symbol sent from the jth transmit antenna, $\mathrm{E}[\tilde{\boldsymbol{s}}] = \boldsymbol{0}$, and $\mathrm{E}[\tilde{\boldsymbol{s}}\tilde{\boldsymbol{s}}^{\mathrm{H}}] = 2\boldsymbol{I}$. The received signal vector $\tilde{\boldsymbol{y}} = [\tilde{y}_1 \cdots \tilde{y}_m] \in \mathbb{C}^m$, where $\tilde{y}_i$ $(i = 1, \ldots, m)$

denotes the signal received at the ith receive antenna, is given by

$$\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{H}}\tilde{\boldsymbol{s}} + \tilde{\boldsymbol{v}}, \tag{1}$$

where

$$\tilde{\boldsymbol{H}} = \begin{bmatrix} \tilde{h}_{1,1} & \cdots & \tilde{h}_{1,n} \\ \vdots & \ddots & \vdots \\ \tilde{h}_{m,1} & \cdots & \tilde{h}_{m,n} \end{bmatrix} \in \mathbb{C}^{m\times n} \tag{2}$$

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 additive white complex Gaussian noise vector with zero mean and covariance matrix of $\sigma_v^2\boldsymbol{I}$. The signal model (1) can be rewritten as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{v}, \tag{3}$$

where

$$\boldsymbol{y} = \begin{bmatrix} \mathrm{Re}\{\tilde{\boldsymbol{y}}\} \\ \mathrm{Im}\{\tilde{\boldsymbol{y}}\} \end{bmatrix}, \ \boldsymbol{H} = \begin{bmatrix} \mathrm{Re}\{\tilde{\boldsymbol{H}}\} & -\mathrm{Im}\{\tilde{\boldsymbol{H}}\} \\ \mathrm{Im}\{\tilde{\boldsymbol{H}}\} & \mathrm{Re}\{\tilde{\boldsymbol{H}}\} \end{bmatrix},$$
$$\boldsymbol{s} = \begin{bmatrix} \mathrm{Re}\{\tilde{\boldsymbol{s}}\} \\ \mathrm{Im}\{\tilde{\boldsymbol{s}}\} \end{bmatrix}, \ \boldsymbol{v} = \begin{bmatrix} \mathrm{Re}\{\tilde{\boldsymbol{v}}\} \\ \mathrm{Im}\{\tilde{\boldsymbol{v}}\} \end{bmatrix}. \tag{4}$$

Since $\tilde{\boldsymbol{s}} \in \{1+\mathrm{j}, -1+\mathrm{j}, -1-\mathrm{j}, 1-\mathrm{j}\}^n$, $\boldsymbol{s}$ is a binary vector whose elements are in $\mathbb{S} = \{1, -1\}$.

III. Conventional Sparse Error Recovery Method

In the conventional sparse error recovery method [5], the non-sparse system model (1), 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 the estimate of $\tilde{\boldsymbol{s}}$ via a detection method, e.g., ZF or MMSE detection. When we use the linear MMSE detection, the estimate is given by

$$\tilde{\boldsymbol{s}}_{\mathrm{est}} = \left(\tilde{\boldsymbol{H}}^{\mathrm{H}}\tilde{\boldsymbol{H}} + \sigma_v^2\boldsymbol{I}\right)^{-1}\tilde{\boldsymbol{H}}^{\mathrm{H}}\tilde{\boldsymbol{y}}. \tag{5}$$

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 $\tilde{\boldsymbol{s}}$. 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 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 (1) as

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

where $\tilde{\boldsymbol{y}}' = \tilde{\boldsymbol{y}} - \tilde{\boldsymbol{H}}\tilde{\boldsymbol{s}}_{\mathrm{est}}^{\mathrm{d}}$. From (7), we can reconstruct the error vector $\tilde{\boldsymbol{e}}$ via compressed sensing algorithms, such as OMP or MMP. Denoting the estimate of the error vector $\tilde{\boldsymbol{e}}$ as $\tilde{\boldsymbol{e}}_{\mathrm{est}}$, we can get the recovered estimate of $\tilde{\boldsymbol{s}}$ as $\tilde{\boldsymbol{s}}_{\mathrm{est}}^{\mathrm{d}} + \tilde{\boldsymbol{e}}_{\mathrm{est}}$.

IV. 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. By applying compressed sensing algorithms to (7), however, we cannot use the discreteness of the error vector and the soft decision of the transmitted signal vector obtained before the error recovery. To achieve a better performance, we propose an error recovery method based on MAP estimation with several relaxations.

A. MAP Estimation

The proposed method uses the real signal model (3) during error recovery. Let $\hat{\boldsymbol{s}} \in \mathbb{R}^{2n}$ be the initial estimate of $\boldsymbol{s}$ and $\hat{\boldsymbol{s}}^{\mathrm{d}} = Q_{\mathbb{S}}(\hat{\boldsymbol{s}}) \in \mathbb{S}^{2n}$, we can transform (3) into

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

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

The MAP estimation of $\boldsymbol{e}$ 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 $\boldsymbol{v}$ has the covariance matrix of $(\sigma_v^2/2)\boldsymbol{I}$, the log likelihood function is given by

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

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 $\boldsymbol{e}$. The objective function to be minimized can be written as

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

B. LLR Calculation

To minimize (10), we need to express $p(e_j)$ explicitly. Since $e_j = s_j - \hat{s}_j^{\mathrm{d}}$, e_j is discrete and take a value only in $\mathbb{B} = \{b_1, b_2, b_3\} = \{-2, 0, 2\}$. The probability $p_{\ell,j} = p(e_j = b_\ell)$ $(\ell = 1, 2, 3)$ is given by

$$p_{1,j} = p(s_j = -1),\ p_{2,j} = p(s_j = +1),\ p_{3,j} = 0 \tag{11}$$

if $\hat{s}_j^{\mathrm{d}} = +1$, and

$$p_{1,j} = 0,\ p_{2,j} = p(s_j = -1),\ p_{3,j} = p(s_j = +1) \tag{12}$$

if $\hat{s}_j^{\mathrm{d}} = -1$. For the calculation of (11) and (12), we use the posterior probabilities $p(s_j = +1 \mid \boldsymbol{y})$ and $p(s_j = -1 \mid \boldsymbol{y})$ instead of the prior probabilities $p(s_j = +1)$ and $p(s_j = -1)$, respectively. To obtain the posterior probability, we calculate the posterior log likelihood ratio (LLR) of the transmitted symbols

$$\lambda_j = \log\frac{p(s_j = +1 \mid \boldsymbol{y})}{p(s_j = -1 \mid \boldsymbol{y})} \tag{13}$$

from the estimate $\hat{\boldsymbol{s}}$. For the reduction of the computational complexity, we assume the independence of each received signal and approximate (13) as

$$\lambda_j = \log \frac{p(\boldsymbol{y} \mid s_j = +1)}{p(\boldsymbol{y} \mid s_j = -1)} \tag{14}$$

$$\approx \sum_{i=1}^{2m} \log \frac{p(y_i \mid s_j = +1)}{p(y_i \mid s_j = -1)}. \tag{15}$$

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 $\sigma^2_{\xi_i^j}$ by using the Gaussian approximation [4], (15) can be rewritten as

$$\sum_{i=1}^{2m} \log \frac{p(y_i \mid s_j = +1)}{p(y_i \mid s_j = -1)} \approx \sum_{i=1}^{2m} \frac{2h_{i,j}\left(y_i - \mu_{\xi_i^j}\right)}{\sigma^2_{\xi_i^j}}. \tag{16}$$

Since $\xi_i^j = \sum_{k=1,k\neq j}^{2n} h_{i,k}s_k + v_i$, $\mu_{\xi_i^j}$ and $\sigma^2_{\xi_i^j}$ are given by

$$\mu_{\xi_i^j} = \sum_{k=1,k\neq j}^{2n} h_{i,k}\mathrm{E}[s_k] \tag{17}$$

$$\sigma^2_{\xi_i^j} = \sum_{k=1,k\neq j}^{2n} h_{i,k}^2 \left(1 - \mathrm{E}[s_k]^2\right) + \frac{\sigma_v^2}{2}. \tag{18}$$

Using the estimates $\hat{s}_1, \ldots, \hat{s}_{2n}$, we obtain $\mu_{\xi_i^j}$ and $\sigma^2_{\xi_i^j}$ by replacing $\mathrm{E}[s_k]$ with

$$\hat{s}'_k = \begin{cases} -1 & (\hat{s}_k < -1) \\ \hat{s}_k & (-1 \leq \hat{s}_k < 1) \\ 1 & (1 \leq \hat{s}_k) \end{cases}, \tag{19}$$

which is bounded in $[-1, 1]$ such that $1 - \mathrm{E}[s_k]^2$ in (18) is positive. From the posterior LLR λ_j obtained in the above manner, the posterior probabilities are given by

$$p(s_j = +1 \mid \boldsymbol{y}) = \frac{e^{\lambda_j}}{1 + e^{\lambda_j}}, p(s_j = -1 \mid \boldsymbol{y}) = \frac{1}{1 + e^{\lambda_j}}. \tag{20}$$

C. Relaxation into Convex Optimization Problem

In the similar way as in [9], we transform the objective function (10) and relax the MAP estimation into a convex optimization problem. The probability $p(e_j)$ can be written as $p(e_j) = \prod_{\ell=1}^{3} p_{\ell,j}^{\delta(b_\ell, e_j)}$, where we define $\delta(\alpha, \beta) = 1$ if $\alpha = \beta$, $\delta(\alpha, \beta) = 0$ if $\alpha \neq \beta$, and $0^0 = 1$. Since the objective function (10) is rewritten as

$$\frac{1}{\sigma_v^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{e}\|_2^2 - \sum_{j=1}^{2n}\sum_{\ell=1}^{3} \delta(b_\ell, e_j) \log p_{\ell,j} \tag{21}$$

$$= \frac{1}{\sigma_v^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{e}\|_2^2 - \sum_{j=1}^{2n}\sum_{\ell=1}^{3} (1 - \|e_j - b_\ell\|_0) \log p_{\ell,j}, \tag{22}$$

we consider the following optimization problem:

$$\underset{\boldsymbol{x}\in\mathbb{B}^{2n}}{\text{minimize}} \;\; \frac{1}{\sigma_v^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \sum_{j=1}^{2n}\sum_{\ell=1}^{3} (\log p_{\ell,j}) \|x_j - b_\ell\|_0. \tag{23}$$

The problem (23) is a combinational optimization problem and is difficult to solve. We thus consider to relax (23) into a convex optimization problem with the idea of compressed sensing. By simply replacing $\mathbb{B}^{2n}$ and ℓ_0 norm with $\mathbb{R}^{2n}$ and ℓ_1 norm respectively, however, the resultant problem cannot be convex because $\log p_{\ell,j}$ will be zero or negative. Hence, we firstly replace $\mathbb{B}^{2n}$ and $\log p_{\ell,j}$ with $\mathbb{R}^{2n}$ and $q_{\ell,j} > 0$ respectively, as

$$\underset{\boldsymbol{x}\in\mathbb{R}^{2n}}{\text{minimize}} \;\; \frac{1}{\sigma_v^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \sum_{j=1}^{2n}\sum_{\ell=1}^{3} q_{\ell,j} \|x_j - b_\ell\|_0, \tag{24}$$

and then relax (24) into the SOAV optimization problem [8] as

$$\underset{\boldsymbol{x}\in\mathbb{R}^{2n}}{\text{minimize}} \;\; \frac{1}{\sigma_v^2}\|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{x}\|_2^2 + \sum_{j=1}^{2n}\sum_{\ell=1}^{3} q_{\ell,j} |x_j - b_\ell|. \tag{25}$$

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

$$\sum_{\ell\in\mathcal{L}_j} (\log p_{\ell,j}) \|x_j - b_\ell\|_0 + C_j = \sum_{\ell\in\mathcal{L}_j} q_{\ell,j} \|x_j - b_\ell\|_0 \tag{26}$$

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 (26). For $\ell \notin \mathcal{L}_j$, $q_{\ell,j}$ is fixed to 0. For example, if $\hat{s}_j^{\mathrm{d}} = 1$, $p_{1,j}, p_{2,j} > 0$ and $p_{3,j} = 0$, then $\mathcal{L}_j = \{1, 2\}$ and the condition (26) becomes $q_{2,j} = \log p_{2,j} + C_j$ and $q_{1,j} = \log p_{1,j} + C_j$. We thus select as $C_j = -\min(\log p_{1,j}, \log p_{2,j}) + \tilde{C}_j$ ($\tilde{C}_j \geq 0$) and obtain $q_{1,j}, q_{2,j} \geq 0$. In this case, $q_{3,j}$ is fixed to 0. In general, the condition (26) results in

$$q_{\ell,j} = \log p_{\ell,j} + \frac{C_j}{|\mathcal{L}_j| - 1} \tag{27}$$

for all $\ell \in \mathcal{L}_j$. Hence, we can obtain a positive $q_{\ell,j}$ by selecting as $C_j = -(|\mathcal{L}_j| - 1)\min_\ell \log p_{\ell,j} + \tilde{C}_j$ ($\tilde{C}_j \geq 0$). It should be noted that in the conventional relaxation method [9], the ℓ_0 norm in the right hand of (26) is replaced with the ℓ_1 norm to keep the value of the objective function in (23) and (25) 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 (25) is convex.

The optimization problem (25) can be solved via proximal splitting methods [10]. Letting $g(\boldsymbol{x}) = \|\boldsymbol{y}' - \boldsymbol{H}\boldsymbol{x}\|_2^2/\sigma_v^2$ and $f_{\boldsymbol{Q}}(\boldsymbol{x}) = \sum_{j=1}^{2n}\sum_{\ell=1}^{3} q_{\ell,j}|x_j - b_\ell|$, we can rewrite (25) as

$$\underset{\boldsymbol{x}\in\mathbb{R}^{2n}}{\text{minimize}} \;\; f_{\boldsymbol{Q}}(\boldsymbol{x}) + g(\boldsymbol{x}). \tag{28}$$

A sequence $\boldsymbol{x}_k$ ($k = 0, 1, \ldots$) converging the solution of (28) can be obtained with the following algorithm:

Algorithm 1. *(Douglas-Rachford Algorithm [10] for (28))*

1) Fix $\varepsilon \in (0, 1), \gamma > 0$, and $\boldsymbol{z}_0 \in \mathbb{R}^{2n}$.

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

$$\begin{cases} \boldsymbol{x}_k = \mathrm{prox}_{\gamma f_{\boldsymbol{Q}}}(\boldsymbol{z}_k) \\ \lambda_k \in [\varepsilon, 2-\varepsilon] \\ \boldsymbol{z}_{k+1} = \boldsymbol{z}_k + \lambda_k(\mathrm{prox}_{\gamma g}(2\boldsymbol{x}_k - \boldsymbol{z}_k) - \boldsymbol{x}_k). \end{cases}$$

For a function $\phi : \mathbb{R}^{2n} \to \mathbb{R}$, its proximity operator $\mathrm{prox}_\phi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is defined as

$$\mathrm{prox}_\phi(\boldsymbol{x}) = \arg\min_{\boldsymbol{u}\in\mathbb{R}^{2n}} \phi(\boldsymbol{u}) + \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{u}\|_2^2. \tag{29}$$

By definition, the proximity operator of $\gamma f_{\boldsymbol{Q}}$ is given by

$$[\mathrm{prox}_{\gamma f_{\boldsymbol{Q}}}(\boldsymbol{x})]_j = \begin{cases} x_j - \gamma(-q_{1,j} - q_{2,j} - q_{3,j}) & (x_j < Q_{1,j}) \\ -2 & (Q_{1,j} \le x_j < Q_{2,j}) \\ x_j - \gamma(q_{1,j} - q_{2,j} - q_{3,j}) & (Q_{2,j} \le x_j < Q_{3,j}) \\ 0 & (Q_{3,j} \le x_j < Q_{4,j}) \\ x_j - \gamma(q_{1,j} + q_{2,j} - q_{3,j}) & (Q_{4,j} \le x_j < Q_{5,j}) \\ 2 & (Q_{5,j} \le x_j < Q_{6,j}) \\ x_j - \gamma(q_{1,j} + q_{2,j} + q_{3,j}) & (Q_{6,j} \le x_j) \end{cases}, \tag{30}$$

where $[\mathrm{prox}_{\gamma f_{\boldsymbol{Q}}}(\boldsymbol{x})]_j$ represents the jth element of $\mathrm{prox}_{\gamma f_{\boldsymbol{Q}}}(\boldsymbol{x})$ and

$$Q_{1,j} = -2 + \gamma(-q_{1,j} - q_{2,j} - q_{3,j}) \tag{31}$$
$$Q_{2,j} = -2 + \gamma(q_{1,j} - q_{2,j} - q_{3,j}) \tag{32}$$
$$Q_{3,j} = \gamma(q_{1,j} - q_{2,j} - q_{3,j}) \tag{33}$$
$$Q_{4,j} = \gamma(q_{1,j} + q_{2,j} - q_{3,j}) \tag{34}$$
$$Q_{5,j} = 2 + \gamma(q_{1,j} + q_{2,j} - q_{3,j}) \tag{35}$$
$$Q_{6,j} = 2 + \gamma(q_{1,j} + q_{2,j} + q_{3,j}). \tag{36}$$

Note that $\mathrm{prox}_{\gamma f_{\boldsymbol{Q}}}(\boldsymbol{x})$ is element-wise function because the minimization in (29) can be performed separately for each element. The proximity operator of $\mathrm{prox}_{\gamma g}$ is given by

$$\mathrm{prox}_{\gamma g}(\boldsymbol{x}) = \left(\boldsymbol{I} + \frac{\gamma}{\sigma_v^2}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}\left(\boldsymbol{x} + \frac{\gamma}{\sigma_v^2}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{y}'\right). \tag{37}$$

By using Algorithm 1, we can solve the optimization problem (25) and obtain the estimate $\hat{\boldsymbol{e}}$ of the error vector $\boldsymbol{e}$. The recovered estimate of the transmitted signal vector $\boldsymbol{s}$ is obtained as $\hat{\boldsymbol{s}}^{\mathrm{d}} + \hat{\boldsymbol{e}}$.

D. 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 initial estimate. The algorithm of the proposed method is summarized as follows:

Algorithm 2. (Error Recovery via SOAV optimization)

1) Get an initial estimate of $\boldsymbol{s}$.
2) Iterate a)–f) for T times.
   a) Calculate the posterior LLR λ_j from the current estimate of $\boldsymbol{s}$.
   b) Compute $p_{\ell,j}$ with (11), (12), and (20).
   c) Relax (23) into (25) by replacing the coefficients $\log p_{\ell,j}$ with $q_{\ell,j}$ satisfying (26).
   d) Fix $\varepsilon \in (0,1), \gamma > 0, K > 0$, and $\boldsymbol{z}_0 \in \mathbb{R}^{2n}$.
   e) For $k = 0, 1, \ldots, K$, iterate

$$\begin{cases} \boldsymbol{x}_k = \mathrm{prox}_{\gamma f_{\boldsymbol{Q}}}(\boldsymbol{z}_k) \\ \lambda_k \in [\varepsilon, 2-\varepsilon] \\ \boldsymbol{z}_{k+1} = \boldsymbol{z}_k + \lambda_k(\mathrm{prox}_{\gamma g}(2\boldsymbol{x}_k - \boldsymbol{z}_k) - \boldsymbol{x}_k) \end{cases}$$

   and let $\hat{\boldsymbol{e}} = \boldsymbol{x}_K$.
   f) Modify the current estimate into $\hat{\boldsymbol{s}}^{\mathrm{d}} + \hat{\boldsymbol{e}}$.
3) Apply $Q_{\mathbb{S}}(\cdot)$ to the current estimate and obtain the final estimate of $\boldsymbol{s}$.

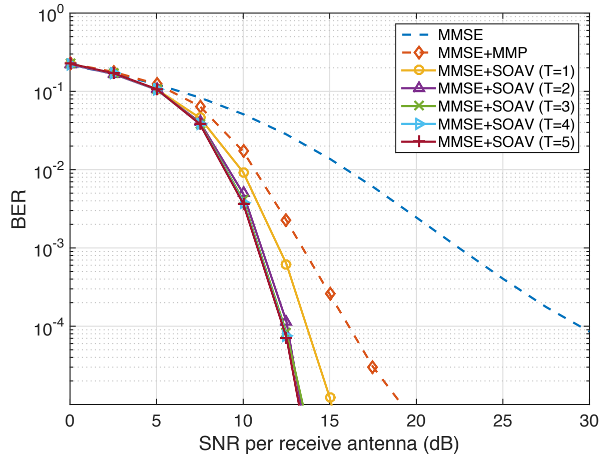

Fig. 1. BER performance for $(n, m) = (32, 32)$

V. Simulation Results

In this section, we evaluate the BER performance of the proposed method and the conventional method 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. 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. For the proposed method, the parameters of the Douglas-Rachford Algorithm are fixed to $K = 30, \varepsilon = 0.1, \gamma = 1, \lambda_k = 1.9$ $(k = 0, 1, \ldots, K)$ and $\boldsymbol{z}_0 = \boldsymbol{0}$. In the proposed relaxation of the MAP estimation, $\tilde{C}_j = 1$ $(j = 1, \ldots, 2n)$ is used.

Figures 1 and 2 show the BER performance for $(n, m) = (32, 32), (128, 128)$, respectively. MMSE, MMSE+MMP, and MMSE+SOAV denote linear MMSE detection, conventional error recovery method via MMP, and our proposed error recovery method via SOAV optimization, respectively. In both recovery methods, the estimate of MMSE detection is used as the initial estimate. The figures show that the proposed method outperforms the conventional method even when $T = 1$, which corresponds to only one iteration of error recovery. It is

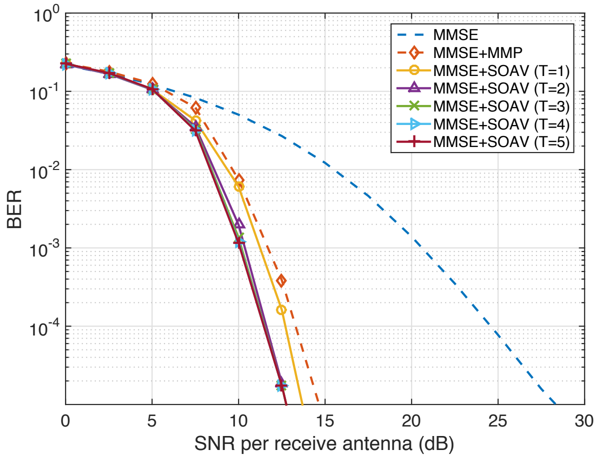

Fig. 2. BER performance for $(n, m) = (128, 128)$

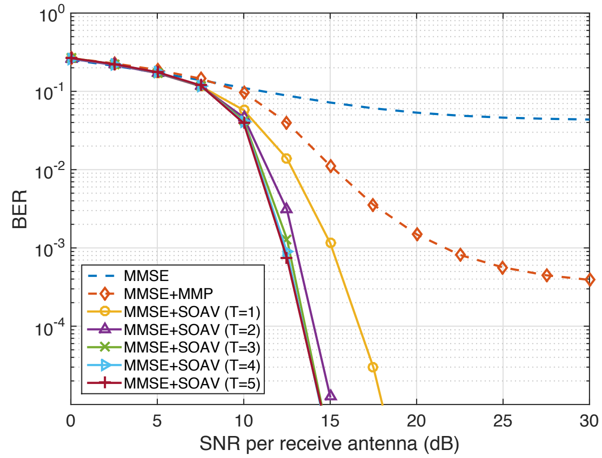

Fig. 4. BER performance for $(n, m) = (128, 96)$

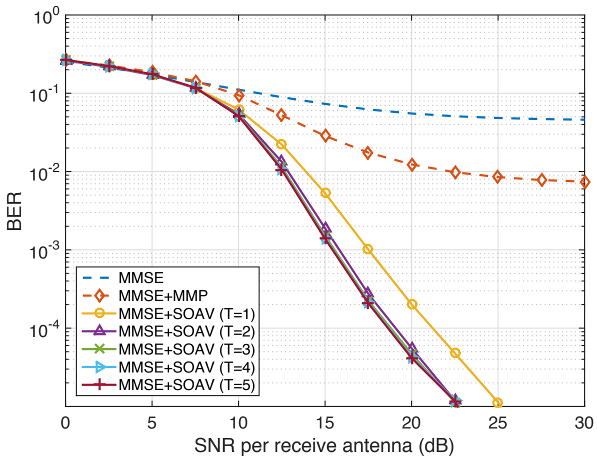

Fig. 3. BER performance for $(n, m) = (32, 24)$

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.

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 [11]. Since the performance of MMSE detection severely degraded in overloaded MIMO, the conventional method also has a poor performance. However, the proposed method performs well even in that case, especially in large-scale systems.

VI. Conclusion

In this paper, we have proposed the MAP estimation-based error recovery for MIMO signal detection, where we can use both the sparsity and the discreteness of the error vector. The proposed method can also use the soft decision obtained by the first detection, while the conventional method uses the hard decision alone. Simulation results show that our proposed method has better BER performance than the conventional method.

Acknowledgment

This work was supported in part by the Grants-in-Aid for Scientific Research no. 15K06064, 15H2252, 15H02668, 15K14006, and 26120521 from the Ministry of Education, Science, Sports, and Culture of Japan and the Telecommunications Advancement Foundation.

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, Apr. 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, Jun. 2007.

[5] J. Choi and B. Shim, "New approach for massive MIMO detection using sparse error recovery," *Proc. IEEE GLOBECOM 2014*, pp. 3754–3759, Dec. 2014.

[6] K. Hayashi, M. Nagahara, 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.

[7] S. Kwon, J. Wang, and B. Shim, "Multipath matching pursuit," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 2986–3001, May 2014.

[8] M. Nagahara, "Discrete signal reconstruction by sum of absolute values," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1575–1579, Oct. 2015.

[9] H. Sasahara, K. Hayashi, and M. Nagahara, "Multiuser detection by MAP estimation with sum-of-absolute-values relaxation," *Proc. IEEE ICC 2016*, May 2016.

[10] 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. Springer New York, vol. 49, pp. 185–212, 2011.

[11] 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.