Massive Overloaded MIMO Signal Detection via Convex Optimization with Proximal Splitting

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

Hampei Sasahara
Engineering School, Tokyo Institute of Technology,
Ookayama, Meguro-ku, Tokyo, 152-8550, Japan
Email: sasahara.h@cyb.mei.titech.ac.jp

Masaaki Nagahara
Institute of Environmental Science and Technology,
The University of Kitakyushu, Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka, 808-0135, Japan
Email: nagahara@ieee.org

Abstract—In this paper, we propose signal detection schemes for massive overloaded multiple-input multiple-output (MIMO) systems, where the number of receive antennas is less than that of transmitted streams. Using the idea of the sum-of-absolute-value (SOAV) optimization, we formulate the signal detection as a convex optimization problem, which can be solved via a fast algorithm based on Douglas-Rachford splitting. To improve the performance, we also propose an iterative approach to solve the optimization problem with weighting parameters update in a cost function. Simulation results show that the proposed scheme can achieve much better bit error rate (BER) performance than conventional schemes, especially in large-scale overloaded MIMO systems.

Index Terms—massive MIMO, overloaded MIMO, proximal splitting methods, Douglas-Rachford algorithm, SOAV optimization

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) systems, where tens or hundreds of antennas are equipped in each transmitter and receiver, are gathering attention as a method to achieve very high spectral efficiency [1], [2]. In massive MIMO systems, low complexity signal detection method is essential because the required computational complexity of MIMO signal detection generally increases along with the increase of the antennas. Some of the candidates for massive MIMO systems are linear signal detections, such as the zero forcing (ZF) and the minimum mean square error (MMSE) detection methods. Besides them, some non-linear detection schemes have also been proposed. The likelihood ascent search (LAS) [3], [4] and the reactive tabu search (RTS) [5], [6] employ local neighborhood search of likelihood and achieve much better performance than linear detection. The graph-based iterative Gaussian detector (GIGD) [7] is well known as a low complexity scheme built upon belief propagation techniques.

In MIMO systems, a sufficient number of receive antennas may not be available because of the limits on the size, weight, cost and/or power consumption of the receiver. Such MIMO systems, where the number of receive antennas are less than that of transmitted streams, are known as overloaded (or underdetermined) MIMO systems. The slab-sphere decoding [8] is a signal detection algorithm based on maximum likelihood (ML) for overloaded MIMO systems to find the solution with lower complexity than that of exhaustive search. Some techniques, such as the pre-voting cancellation [9] and the virtual channel [10], transform overloaded systems into non-overloaded systems to apply conventional MIMO signal detection. [11] and [12] employ the ideas in [8] and [9] to achieve a good performance with lower complexity. For massive overloaded MIMO systems, however, these schemes are not practical because their complexity is still high, while the performance of low-complexity detection for massive MIMO systems is considerably degraded in the overloaded scenario. To further reduce the complexity, the enhanced reactive tabu search (ERTS), which is an extension of the random restart reactive tabu search (RTS) [6], has been recently proposed [13]. ERTS employs RTS iteratively while varying the initial point of the search randomly until the estimate by RTS satisfies a certain condition. In [13], it is shown that ERTS can achieve a comparable performance to the optimal ML detection with affordable computational complexity for overloaded MIMO systems with tens of antennas. With hundreds of antennas, however, ERTS requires prohibitive computational complexity to achieve such performance because the required number of RTSs increases with the number of antennas.

In this paper, we propose a massive overloaded MIMO signal detection scheme with much lower complexity than that of conventional schemes [8]–[12]. We formulate the signal detection problem as a convex optimization problem, where the idea is based on the sum-of-absolute-value (SOAV) optimization [14], which is a technique to reconstruct a discrete-valued vector from its linear measurements. The optimization problem can be efficiently solved with proximal splitting methods [15] even for underdetermined systems. To improve the performance, we extend SOAV optimization to weighted-SOAV optimization, where the prior information about the discrete-valued vector can be used, and propose an iterative
approach, named iterative weighted-SOA (IW-SOA), using the estimate in the previous iteration as the prior information. Since the weighted-SOA optimization problem can also be efficiently solved with proximal splitting methods, IW-SOA can detect the transmitted signals with low computational complexity. Simulation results show that IW-SOA can achieve much better bit error rate (BER) performance than conventional signal detection schemes especially in large-scale overloaded MIMO systems.

In the rest of the paper, we use the following notations. Superscript $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and the Hermitian transpose, respectively. Mathematica symbols, $j$, $I$, and $0$ represent the imaginary unit, the identity matrix, a vector whose elements are all 1, and a vector whose elements are all 0. For a vector $a = [a_1 \ldots a_N]^T \in \mathbb{R}^N$, we define the $\ell_1$ and $\ell_2$ norms of $a$ as

$$
\|a\|_1 = \sum_{i=1}^N |a_i| \quad \text{and} \quad \|a\|_2 = \sqrt{\sum_{i=1}^N a_i^2},
$$

respectively. $\Pr(A)$ denotes the probability of an event $A$ and $E[\cdot]$ stands for the expectation operator.

II. SYSTEM MODEL

Here 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 the quadrature phase shift keying (QPSK) and define the alphabet of the transmitted symbol as $S = \{1+j, -1+j, -1-j, 1-j\}$. The transmitted signal vector $s = [s_1 \ldots s_n]^T \in \mathbb{R}^n$ is composed of signals transmitted from $n$ transmit antennas, where $s_j (j = 1, \ldots, n)$ denotes the symbol sent from the $j$th transmit antenna, $E[s] = 0$, and $E[ss^H] = 2I$. The received signal vector $y = [y_1 \ldots y_m] \in \mathbb{C}^m$, where $y_i \ (i = 1, \ldots, m)$ denotes the signal received at the $i$th receive antenna, is given by

$$
y = Hs + v, \quad (2)
$$

where

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

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

$$
y = Hs + v, \quad (4)
$$

where

$$
s = \begin{bmatrix}
\text{Re}(s) \\
\text{Im}(s)
\end{bmatrix}, \quad v = \begin{bmatrix}
\text{Re}(v) \\
\text{Im}(v)
\end{bmatrix},
$$

Since $s \in \{1+j, -1+j, -1-j, 1-j\}^n$, $s$ is a binary vector whose elements are 1 or -1.

III. PROPOSED SIGNAL DETECTION SCHEMES

In this section, we propose signal detection schemes based on SOAV optimization for massive overloaded MIMO systems. We briefly review SOAV optimization in Sect. III-A and propose a signal detection scheme in Sect. III-B. In Sect. III-C, we also propose an iterative approach to improve the performance.

A. SOAV Optimization

SOAV optimization [14] is a technique to reconstruct an unknown discrete-valued vector as $x = [x_1 \cdots x_N]^T \in \{c_1, \ldots, c_P\}^N \subset \mathbb{R}^N$ from its linear measurements $\eta = Ax$, where $A \in \mathbb{R}^{M \times N}$. If we assume $\Pr(x_i = c_p) = 1/P \ (p = 1, \ldots, P)$ for all $x_i \ (i = 1, \ldots, N)$, each of $x - c_1, \ldots, x - c_P$ has approximately $N/P$ zero elements. Based on this property and the idea of $\ell_1$ optimization in compressed sensing [16], SOAV optimization solves

$$
\min_{x \in \mathbb{R}^N} \frac{1}{P} \sum_{p=1}^P \|x - c_p\|_1
$$

subject to $\eta = Ax \quad (6)$

to reconstruct $x$ from $\eta$.

B. Proposed Signal Detection via SOAV Optimization

In MIMO systems, the transmitted signal vector $s$ is commonly discrete and the received signal vector $y$ can be regarded as its linear observations if the noise can be ignored. Since each element of $s$ is 1 or -1 for the case with QPSK, we can formulate the signal detection problem as SOAV optimization, i.e.,

$$
\min_{x \in \mathbb{R}^{2n}} \frac{1}{2} \|z - 1\|_1 + \frac{1}{2} \|z + 1\|_1
$$

subject to $y = Hz \quad (7)$

Since the received signal vector $y$ contains the additive noise as in (4), we modify the optimization problem as follows:

$$
\min_{x \in \mathbb{R}^{2n}} \frac{1}{2} \|z - 1\|_1 + \frac{1}{2} \|z + 1\|_1 + \frac{\alpha}{2} \|y - Hz\|_2^2 \quad (8)
$$

by using the idea of $\ell_1$-$\ell_2$ optimization. Here, $\alpha > 0$ is a given constant. The solution of (8) can be obtained with the following theorem [15].

**Theorem 1.** Let $\phi_1, \phi_2 : \mathbb{R}^{2n} \to (-\infty, \infty]$ be lower semicontinuous convex functions and $(\text{ri dom } \phi_1) \cap (\text{ri dom } \phi_2) \neq \emptyset$. In addition, $\phi_1(z) + \phi_2(z) \to \infty \text{ as } \|z\|_2 \to \infty$ is assumed. A sequence $z_k \ (k = 0, 1, \ldots)$ converging to the solution of

$$
\min_{z \in \mathbb{R}^{2n}} \phi_1(z) + \phi_2(z) \quad (9)
$$
where $f$ represents $\mathbb{R}^{2n} \to \mathbb{R}$ is defined as

$$\text{prox}_\phi(z) = \arg \min_{w \in \mathbb{R}^{2n}} \phi(u) + \frac{1}{2} \|z - u\|^2. \tag{10}$$

Algorithm 1. (Douglas-Rachford Algorithm)

1) Fix $\varepsilon \in (0, 1)$, $\gamma > 0$, and $r_0 \in \mathbb{R}^{2n}$.
2) For $k = 0, 1, 2, \ldots$, iterate

$$\begin{align*}
\mathbf{z}_k &= \text{prox}_{\gamma \phi_2}(r_k) \\
\lambda_k &= \left\lfloor \varepsilon, 2 - \varepsilon \right\rfloor \\
\mathbf{r}_{k+1} &= \mathbf{r}_k + \lambda_k (\text{prox}_{\gamma \phi_1}(2\mathbf{z}_k - \mathbf{r}_k) - \mathbf{z}_k).
\end{align*}$$

We can rewrite (8) as

$$\min_{z \in \mathbb{R}^{2n}} f(z) + g(z), \tag{11}$$

where $f(z) = \|z - 1\|_2^2/2 + \|z + 1\|_2^2/2$ and $g(z) = \alpha \|y - Hz\|_2^2/2$. The proximity operators of $\gamma f(z)$ and $\gamma g(z)$ can be obtained as

$$\begin{align*}
[\text{prox}_{\gamma f}(z)]_j &= \begin{cases} 
\gamma_z & (z_j < -1 - \gamma) \\
-1 & (-1 - \gamma \leq z_j < -1) \\
z_j & (-1 \leq z_j \leq 1) \\
1 & (1 \leq z_j < 1 + \gamma) \\
z_j - \gamma & (1 + \gamma \leq z_j)
\end{cases} \\
\text{prox}_{\gamma g}(z) &= (I + \alpha \gamma H^T H)^{-1}(z + \alpha \gamma H^T y), \tag{13}
\end{align*}$$

and

$$\begin{align*}
[\text{prox}_{\gamma f}(z)]_j &= \begin{cases} 
\gamma_z & (z_j < -1 - \gamma) \\
-1 & (-1 - \gamma \leq z_j < -1) \\
z_j & (-1 \leq z_j \leq 1) \\
1 & (1 \leq z_j < 1 + \gamma) \\
z_j - \gamma & (1 + \gamma \leq z_j)
\end{cases} \\
\text{prox}_{\gamma g}(z) &= (I + \alpha \gamma H^T H)^{-1}(z + \alpha \gamma H^T y), \tag{13}
\end{align*}$$

respectively, where $[\text{prox}_{\gamma f}(z)]_j$ ($j = 1, \ldots, 2n$) represents the jth element of $\text{prox}_{\gamma f}(z)$. Note that $[\text{prox}_{\gamma f}(z)]_j$ is a function of $z_j$ only as shown in Fig. 1. By solving (8) with the Douglas-Rachford algorithm, the estimate of the transmitted signal vector $s$ can be obtained.

### C. Proposed Iterative Approach, IW-SOA

Assuming that we have information on prior probabilities of $w^+_j = \Pr(s_j = 1)$ and $w^-_j = \Pr(s_j = -1)$, we extend the problem of (8) to weighted-SOA optimization problem as

$$\min_{z \in \mathbb{R}^{2n}} \sum_{j=1}^{2n} \left( w^+_j |z_j - 1| + w^-_j |z_j + 1| \right) + \frac{\alpha}{2} \|y - Hz\|^2. \tag{14}$$

If there is no prior information about $s$, i.e., $w^+_j = w^-_j = 1/2$, the optimization problem (14) is equivalent to (8). If $w^+_j > w^-_j$ then $\arg \min_{z_j} f_{w_j}(z_j) = 1$, where $f_{w_j}(z_j) = w^+_j |z_j - 1| + w^-_j |z_j + 1|$, thus the solution of $z_j$ in (14) tends to take the value close to 1, and vice versa. The optimization problem (14) can also be solved by using the Douglas-Rachford algorithm.

The proximity operator of

$$\gamma f_w(z) = \gamma \sum_{j=1}^{2n} \left( w^+_j |z_j - 1| + w^-_j |z_j + 1| \right) \tag{15}$$

can be written as

$$[\text{prox}_{\gamma f_w}(z)]_j = \begin{cases} 
z_j + \gamma & (z_j < -1 - \gamma) \\
-1 & (-1 - \gamma \leq z_j < -1) \\
z_j & (-1 \leq z_j \leq 1) \\
1 & (1 \leq z_j < 1 + \gamma) \\
z_j - \gamma & (1 + \gamma \leq z_j)
\end{cases}, \tag{16}$$

as shown in Fig. 2, where $d_j = w^+_j - w^-_j$. By solving the optimization problem (14) via the Douglas-Rachford algorithm with $\text{prox}_{f_w}$ and $\text{prox}_{g}$, a new estimate of the transmitted signal vector $s$ can be obtained.

The prior information on $s$ is not available in a common scenario, however, assuming iterative approach, the estimate in the previous iteration can be used to obtain the prior probabilities. Specifically, in the proposed iterative approach
schemes are set as $\hat{s}_j = \begin{cases} 0 & (\hat{s}_j < -1) \\
1 + \hat{s}_j & (-1 \leq \hat{s}_j < 1) \\
1 & (1 \leq \hat{s}_j) \end{cases}$ (17) and

$$w_j^+ = \left\{ \begin{array}{ll}
1 & (\hat{s}_j < -1) \\
1 - \hat{s}_j & (-1 \leq \hat{s}_j < 1) \\
0 & (1 \leq \hat{s}_j)
\end{array} \right. \tag{18}$$

where $\hat{s}_j$ is the estimate of $s_j$ in the previous iteration. Fig. 3 shows $w_j^+$ and $w_j^-$ as a function of $\hat{s}_j$. $w_j^+$ is large when $\hat{s}_j$ is large, and $w_j^-$ is large when $\hat{s}_j$ is small. This is because the estimates close to 1 or −1 will be more reliable than those close to 0. The proposed algorithm of IW-SOA is summarized as follows:

Algorithm 2. (Proposed Signal Detection via IW-SOAV)

1. Let $\hat{s} = \mathbf{0}$ and iterate a)–c) for $L$ times.
   a) Compute $w_j^+$, $w_j^-$ with (17),(18).
   b) Fix $\varepsilon \in (0,1), \gamma > 0, K > 0$, and $r_0 \in \mathbb{R}^{2n}$.
   c) For $k = 0, 1, 2, \ldots, K$, iterate
   $$\begin{align*}
   z_k &= \text{prox}_{\gamma \beta}(r_k) \\
   \lambda_k &= [\varepsilon, 2 - \varepsilon] \\
   r_{k+1} &= r_k + \lambda_k (\text{prox}_{\gamma \beta}(2z_k - r_k) - z_k)
   \end{align*}$$
   and let $\hat{s} = z_K$.
2. Obtain $\hat{s}$ as the final estimate of $s$.

IV. SIMULATION RESULTS

In this section, we evaluate the BER performance of the proposed scheme by computer simulation comparing with that of conventional detection methods. In the simulation, flat Rayleigh fading channels are assumed and $\mathbf{H}$ is composed of independent and identically distributed complex Gaussian random variables with zero mean and unit variance. The parameter $\alpha$ in (14) is selected as $\alpha = 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 1$, and 1 for SNR per receive antenna of 0, 5, 10, 15, 20, 25, and 30 (dB), respectively. The other parameters of the proposed schemes are set as $K = 50, \varepsilon = 0.1, \gamma = 1, \lambda_k = 1.9$ ($k = 0, 1, \ldots, K$), and $r_0 = \mathbf{0}$.

Figs. 4–6 shows the BER performance for overloaded MIMO systems with $(n, m) = (25, 16), (150, 96)$, and $(200, 128)$, respectively, where the ratio $m/n$ is fixed to be 0.64 for all cases. In the figures, MMSE, GIGD, and ERTS represent the linear MMSE detection, the belief propagation-based detection [7], and the massive overloaded MIMO signal detection proposed in [13], respectively. The parameters of ERTS are the same as those in [13], e.g., the maximum number of RTTs is $N_{\text{RTS}} = 500$ and the maximum number of iterations in RTS is $N_{\text{itr}} = 300$. IW-SOA denotes our proposed scheme shown in Algorithm 2. Comparing the performance of IW-SOA with $L = 1$ and 10 in the figures, where $L$ is the number of iterations, we can see that the BER performance is improved with the proposed iterative approach. Although the performance of ERTS is much better than that of IW-SOA in Fig. 4, it considerably degrades for larger number of antennas as shown in Fig. 5 and 6. This is because, 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 RTTs, the computational complexity could be prohibitive to achieve comparable performance as IW-SOA. Specifically, given that the computational complexity of ERTS is $\mathcal{O}(n^3) + \mathcal{O}(N_{\text{RTS}}n^2)$ in the worst case, and the number of all candidates of the transmit signal vector increases exponentially with the number of transmit antennas, the required $N_{\text{RTS}}$ to keep good performance will increase more rapidly than $n$. On the other hand, the computational complexity of IW-SOA is $\mathcal{O}(n^3)$, which is dominated by
the calculation of $(I + \alpha \gamma H^T H)^{-1}$ in (13). Note that the calculation of $(I + \alpha \gamma H^T H)^{-1}$ is required only once, and thus the corresponding computational cost does not grow with $K$ or $L$. If the ratio $m/n$ is fixed, the computational complexity of IW-SOAV is the same order as that of MMSE detection $O(m^3) + O(mn)$ for an overloaded scenario. From the figures, our proposed scheme has much better performance than the other schemes for the large-scale overloaded MIMO systems.

Fig. 7 shows the BER performance versus the number of receive antennas $m$ for $n = 150$ and SNR per receive antenna of 20 dB. We can observe that IW-SOAV with $L = 10$ requires less antennas than other schemes to achieve good BER performance. For $BER = 10^{-4}$, IW-SOAV can reduce approximately 10 receive antennas compared to the conventional ERTS.

V. CONCLUSION

In this paper, we have proposed a massive overloaded MIMO signal detection scheme, namely IW-SOAV, which iteratively solves the weighted-SOA V optimization problem while updating its parameters. Simulation results show that IW-SOAV can achieve much better performance than conventional massive MIMO detection schemes, especially in large-scale overloaded MIMO systems. Future work includes the integration of our proposed scheme and soft channel decoding schemes, such as low density parity check (LDPC) codes and turbo codes.

REFERENCES