Article:

Zhang, X., Zhou, Z., Zhang, L. orcid.org/0000-0002-4535-3200 et al. (2 more authors) (2025) Multiuser Detection with Compressive Sensing Iterative Reweighed Approach for Grant-Free MIMO-NOMA Systems. IEEE Transactions on Vehicular Technology, 74 (1). 1742 -1746. ISSN 0018-9545

https://doi.org/10.1109/TVT.2024.3455188

Multiuser Detection with Compressive Sensing Iterative Reweighed Approach for Grant-Free MIMO-NOMA Systems

Xiaoxu Zhang, *Member, IEEE*, Ziyang Zhou, Li Zhang, *Senior Member, IEEE*, Pingzhi Fan, *Fellow, IEEE*, and Zheng Ma, *Member, IEEE*

***Abstract*—In this paper, we focus on the efficient multiuser detection (MUD) problem for multiple-input multiple-output (MIMO) enabled grant-free non-orthogonal multiple access (GF-NOMA) system for massive machine-type communications (MTC). The inherent sparsity of mMTC motivates us to make use of compressive sensing (CS) technology to address the MUD problem. This paper discusses the use of an iterative reweighed (IR) scheme combined with the majorization-minimization (MM) algorithm to recover sparse signals under the Grant-free MIMO-NOMA model. Numerical and simulation results demonstrate that Grant-Free MIMO-NOMA with the proposed IR scheme is capable of reconstructing sparse signals with unknown user activity factors and the proposed algorithm outperforms conventional multiuser detectors.**

***Index Terms*—Massive machine-type communications, Grant-Free MIMO-NOMA, multiuser detection, iterative reweighed.**

I. Introduction

The sixth generation (6G) of communication promises to revolutionise connectivity and drive unprecedented advancements across numerous industries. A significant feature of 6G is massive machine-type communications (MTC), which seeks to facilitate an enormous quantity of connected devices such as Internet of Things (IoT) devices [1], sensors, and smart devices. Furthermore, the forthcoming 6G network will utilise multiple-input multiple-output (MIMO) technology to enhance data transmission and augment network capacity.

In decades of research, MIMO technology has mainly focused on point-to-point MIMO systems with multiple transmitter and receiver antennas. For the multiuser-MIMO (MU-MIMO) system, the base station (BS) uses more sophisticated equipment with multiple antennas to transmit signals to the user terminal equipped with a single antenna. Multi-antenna deployment of MIMO also brings significant power consumption and implementation complexity to the radio freqency (RF) of the antennas. In [2], Z. Lin et al. investigate three different hybrid beam forming (BF) architectures, resolved the excessive RF power in multiple antennas, making multiple antennas more practical. The grant-free non-orthogonal multiple access (GF-NOMA) technique has been widely studied recently as it eliminates the need for interactive handshaking process and improves the spectral efficiency. The Grant-Free process greatly reduce transmission latency and signaling overhead [4]. This technique can be applied to MU-MIMO system, and form a Grant-Free MIMO-NOMA system. However, the complexity of data transmission and processing increases significantly due to the increase of large number of antennas and it is necessary to perform accurate multiuser detection (MUD) for the Grant-Free access. Therefore, there is an urgent need for efficient and accurate MUD algorithms of Grant-Free MIMO-NOMA system.

Since there are a large number of users in mMTC and very few active users, we consider that the user's transmitted signal is sparse. Compressive sensing (CS) reduces the transmission and processing burden by exploiting the sparsity of the signal. Therefore, there have been many researches on CS to solve the MUD problem [6]-[11]. The author in [6] propose a FOCUSS (FOCal Underdetermined System Solver) scheme for GSM-MIMO system. H. Iimori [7] derived new Bayesian message passing rules based on Gaussian approximation, which enables a novel joint activity, channel and data estimation. In [8], F. Li proposed adopting sparse Bayesian learning (SBL), a coupled hierarchical Gaussian framework, to reduce pilot pollution caused by imperfect channel estimation. However, it cannot overcome the computational complexity associated with Bayesian algorithms. X. Zhang [9] proposed SBL and PCSBL under single-slot and multi-slot, and in [10] [11], H. Zhu and S. Adnan use traditional MUD and CS greedy algorithms respectively. However, while the complexity has been reduced, it is also very important to enhance recovery performance.

The previous study of Grant-Free MIMO-NOMA lacked MUD algorithms designed from the perspective of user sparsity. Therefore, the dimensionality of the processed signals is large. Since there are a large number of users, but only a very small number of them are active, the transmitted MIMO signal has a sparse feature. Taking advantage of this property, this paper adopts the CS method to perform MUD

This work was partially funded by the National Science Foundation of China Project (No. 62020106001), and in part by the Sichuan Science and Technology Program (No. 2022NSFSC0910). The work of Zheng Ma was supported by NSFC project under Grant U2268201, 62271419 and 62361136810.

X. Zhang, Z. Zhou, P. Fan, and Z. Ma are with Southwest Jiaotong University, Chengdu 611756, China (e-mail: xiaoxuzhang@swjtu.edu.cn, zhouziyang@my.swjtu.edu.cn, p.fan@ieee.org, zma@home.swjtu.edu.cn).

L. Zhang is with School of Electronic and Electrical Engineering, University of Leeds, Leeds, LS2 9JT, U.K (e-mail: L.X.Zhang@leeds.ac.uk).

of the Grant-Free MIMO-NOMA uplink system. Previously, X. Zhang has verified the effectiveness of the iterative reweighting algorithm for single-measurement vector models and multi-measurement vector models in [12] and [13] respectively, which do not consider the use of MIMO systems to improve the transmission efficiency. In this paper, we extend the iterative algorithm to multi-antenna scenarios and verify the effective performance of the IR scheme under the Grant-Free MIMO-NOMA system.

The main contributions of this paper are summarized as follows:

- We convert the problem of signal recovery in multi-antenna scenarios into a CS problem in Grant-Free MIMO-NOMA systems and enhance the traditional CS model. More precisely, we replace the $l1$-norm with the Log-Sum and the exponential functions.
- We combine majorization-minimization (MM) algorithm with IR algorithm, and propose an efficient MUD algorithm. The proposed MIMO-IR scheme does not require a prior information about the user activity factor, which is more realistic than the general MUD approaches.
- Simulation proves that the MIMO-IR algorithm has better recovery performance and lower computational complexity compared with other MUD algorithms in the MIMO system.

This paper is organized as follows: In Section II, we describe the Grant-Free MIMO-NOMA system. In Section III, we derive two different MIMO-IR schemes respectively. In Section IV, we present the simulation results and analysis to demonstrate the reliability and feasibility of the algorithm. In Section V, we conclude the paper.

II. System Model

The multiplexing gain of point-to-point MIMO is dependent on a favorable propagation environment and high signal-to-noise ratio. In other words, point-to-point MIMO performance degrades significantly under line-of-sight (LOS) propagation or when the terminal is located at the edge of the cell [3]. MU-MIMO can effectively overcome these problems by sharing antenna resources and space resources among multiple users. In addition, MU-MIMO can significantly reduce the interference among users at the edge of the cell through inter-user co-operation and interference cancellation techniques, thus improving the communication performance of these users. According to recent research [5], it has been found that in large-scale M2M systems, with a sufficiently large number of antennas, the channel vectors will eventually become asymptotically orthogonal to each other, i.e.

$$(1/M)\boldsymbol{h}_i^H\boldsymbol{h}_j = 0, \quad i \neq j. \tag{1}$$

where $\boldsymbol{h}_i$ and $\boldsymbol{h}_j$ represent the channel vector of user i and j. However, traditional MIMO systems' capacity is limited by the number of users and antennas. GF-NOMA allows multiple users to share the same time, frequency, and spatial resources. This overcomes the capacity limitations of conventional MIMO systems and enables users with poorer channel conditions to achieve good communication performance through non-orthogonal resource allocation. Compared to other multiple access schemes, NOMA provides better spectral efficiency and user fairness.

We consider a uplink single cell Grant-Free MIMO-NOMA system with K single-antenna users and a BS with M antennas. In this MU-MIMO system, a BS can simultaneously serve K independent terminals. For a transmission of K users, the received signal can be formulated as follows:

$$\mathbf{Y} = \mathbf{HX} + \mathbf{W}, \tag{2}$$

where $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_p) \in \mathbb{C}^{M\times P}$ is the received signal of the M antennas. It is equivalent to a narrow band system when the number of subcarriers $N = 1$, and $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_p) \in \mathbb{C}^{K\times P}$ is the transmitted signal with length P of the K users at the transmit antennas. $\mathbf{x}_p \in \mathbb{C}^{NK\times 1}$ is the pth symbol sequence of K users. When the number of subcarriers $N > 1$, $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_p) \in \mathbb{C}^{NK\times P}$, the row elements of each element $\mathbf{x}_p$ are expanded into $N \times 1$ vectors and stacked up sequentially. If the kth device is active, the device transmits a Binary Phase Shift Keying (BPSK) modulated signal of length P, $\mathbf{x} \in \mathcal{A} = \{-1, +1\}$, and the inactive devices remain 0. $\mathbf{W} = (\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_p) \in \mathbb{C}^{M\times P}$ is a zero-mean circularly symmetric complex Gaussian matrix with a unit covariance matrix of M receive antenna of each symbols. According to the power allocation strategy for power-domain NOMA, in order to achieve fairness and maximize system capacity, signals with higher power are allocated to users with poorer channel conditions, and signals with lower power are allocated to users with better channel conditions. We denote the channel power allocated by NOMA by diagonal matrix $\xi \in \mathbb{C}^{K\times K}$. $\mathbf{H} \in \mathbb{C}^{M\times K}$ is a complex channel matrix, which is the product of complex small-scale fading, large-scale fading coefficients [3] and the power allocation weights:

$$\mathbf{H} = \mathbf{G}\mathbf{S}_\gamma^{1/2}\xi, \tag{3}$$

where $M \times K$ matrix $\mathbf{G}$ accounts for small scale fading coefficients, the diagonal elements of $K \times K$ diagonal matrix $\mathbf{S}_\gamma^{1/2}$ is the terminal's large-scale fading coefficient and ξ is the power allocation weights.

III. Proposed MIMO-IR scheme

In this section, we present the iterative reweighted based CS algorithm for MIMO-NOMA systems. The task is to obtain the estimated signal in (1). In order to recover the sparse signal in $\mathbf{x}_p$, we describe the signal to be detected $\mathbf{y}_p$ by using the minimum atoms in CS. Therefore, the problem can be formulated as follows:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \|\mathbf{x}_p\|_0 \\ \text{s.t.} \quad & \mathbf{y}_p = \mathbf{H}\mathbf{x_p}, \end{aligned} \tag{4}$$

where $\|\mathbf{x}_p\|_0$ represents the sparsity of $\mathbf{x}_p$. However, the traditional CS $l0$-norm problem is an NP-Hard problem, which is generally transformed into $l1$-norm [14] or $l2$-norm [15] in previous research.

In this paper, we consider sparse signal recovery using two different forms of the objective function, which has been shown to outperform $l1$-norm or $l2$-norm in terms of performance as a signal recovery function [12].

A. Exponential form

To facilitate the solution of the CS problem and to ensure convergence, it is advisable to formulate a convex objective function. This choice guarantees the existence and uniqueness of a globally optimal solution and often allows the use of efficient algorithms. Therefore, we aim to reformulate the exponential function e^x into a convex form. Substituting the above formula $l0$-norm with exponential form, we can get:

$$\begin{aligned} \min_{\boldsymbol{x}} F(\boldsymbol{x}) &= \sum_{k=1}^{K}(1-\exp\left(||x_k||^2/\epsilon\right)) \\ \text{s.t.} \quad & ||\mathbf{y}_p-\mathbf{H}\mathbf{x}_p||_2 \leq \eta, \end{aligned} \tag{5}$$

in which ϵ is a regularisation constant and $\epsilon > 0$. η is a tolerance parameter for noise, which usually takes on a very small value. x_k is the kth element of the vector $\mathbf{x}_p$. The essence of MIMO multi-antenna extends the problem to the matrix side, so the above constraints can be transformed into:

$$\begin{aligned} \min_{\boldsymbol{X}} F(\boldsymbol{X}) &= \sum_{k=1}^{K}(1-\exp\left(||\boldsymbol{X_k}||^2/\epsilon\right)) \\ \text{s.t.} \quad & ||\mathbf{Y}-\mathbf{H}\mathbf{X}||_F \leq \eta, \end{aligned} \tag{6}$$

where $\boldsymbol{X_k}$ is the kth column of the matrix $\mathbf{X}$. In order to improve computational efficiency and find optimal solutions, an unconstrained optimisation problem can be formulated by adding a penalty term to the objective function. The resulting Lagrange equation is expressed as follows:

$$\min_{\mathbf{X}} L(\mathbf{X}) \triangleq \sum_{k=1}^{K}(1-\exp\left(||\boldsymbol{X_k}||^2/\epsilon\right)) + \lambda\left\|\mathbf{Y}-\mathbf{H}\mathbf{X}\right\|_F^2, \tag{7}$$

where $\lambda > 0$ balances data fit and sparsity. In MM algorithms, the surrogate function simplifies optimization by constructing an auxiliary function, aiding convergence in each iterative step. It's crucial for the surrogate function to be differentiable and convex to ensure progressive convergence towards stability:

$$\begin{aligned} Q(\mathbf{X}|\tilde{\mathbf{X}}^{(t)}) \triangleq \sum_{k=1}^{K}\Big(&1-e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon}+ \\ & e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon}(|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon-|\sqrt{P}x_{kp}|^2/\epsilon)\Big), \end{aligned} \tag{8}$$

where $\widetilde{x}_{kp}^{(t)}$ is the kth users in pth transmit symbols at tth iteration.

In order to prove that the above surrogate function is the upper bound on the original objective function, i.e., $J(\mathbf{X}) = Q(\mathbf{X}|\widetilde{\mathbf{X}}^{(t)}) - F(\mathbf{X}) \geq 0$, the proof is as follows.

Proof: Notice that the values of the column squares of the same row element in $\mathbf{X}$ are the same:

$$\begin{aligned} \sum_{k=1}^{K}(1-e^{(||\boldsymbol{X_k}||^2/\epsilon)}) &= \sum_{k=1}^{K}(1-e^{((x_{k1}^2+x_{k2}^2+\ldots+x_{kp}^2)/\epsilon)}) \\ &= \sum_{k=1}^{K} 1-e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/}\epsilon. \end{aligned} \tag{9}$$

The upper bound we need to prove is as follows:

$$\begin{aligned} &\sum_{k=1}^{K}(1-e^{(||\boldsymbol{X_k}||^2/\epsilon)}) \\ &\leq \sum_{k=1}^{K}\Big(1-e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon}+ \\ &\quad e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon}(|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon-|\sqrt{P}x_{kp}|^2/\epsilon)\Big). \end{aligned} \tag{10}$$

Then the difference between the surrogate function and the original objective function $J(\mathbf{X})$ is as follows:

$$\begin{aligned} J(\mathbf{X}) = &\sum_{k=1}^{K}\Big(1-e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon}-(1-e^{|\sqrt{P}x_{kp}|^2/\epsilon}) \\ &+e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon}(|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon-|\sqrt{P}x_{kp}|^2/\epsilon)\Big). \end{aligned} \tag{11}$$

To facilitate the calculation, we define the following parameters:

$$\begin{aligned} \alpha &= |\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon \\ \beta &= |\sqrt{P}x_{kp}|^2/\epsilon. \end{aligned} \tag{12}$$

Substituting the parameters in (12) into $J(\mathbf{X})$ results in:

$$\begin{aligned} J(\mathbf{X}) &= \sum_{k=1}^{K}\big(1-e^a-(1-e^b) \\ &\quad +e^a(a-b)\big) \\ &= \sum_{k=1}^{K}\left(e^b-e^a+e^a(a-b)\right) \\ &= \sum_{k=1}^{K} e^a\left((e^{b-a}-1)-(b-a)\right) \\ &= \sum_{k=1}^{K} e^a f(x). \end{aligned} \tag{13}$$

Notice that:

$$e^a = e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon} \geq 0, \tag{14}$$

we set $b-a=x$, then the part of (13):

$$f(x) = e^x - 1 - x, \tag{15}$$

and

$$f'(x) = e^x - 1, \tag{16}$$

when $x=0$, we have $f(0)=0$. And when $x \geq 0$, $f'(x) \geq 0$; when $x \leq 0$, $f'(x) \leq 0$. We conclude that $f(0)$

is the minimum value, and $f(x) \geq 0$, then $J(\mathbf{X}) \geq 0$.
End of proof.

In summary, the upper bound is reached at $\alpha = \beta$, i.e., $\widetilde{x}_{kp}^{(t)} = x_{kp}$, and the final optimal formulation expressed as:

$$\begin{aligned}
&\min_{\mathbf{X_p}} Q(\mathbf{X}|\tilde{\mathbf{X}}^{(t)}) + \lambda \left\|\mathbf{Y} - \mathbf{H}\mathbf{X}\right\|_F^2 \\
&= \min_{\mathbf{X_p}} \sum_{k=1}^{K} \Big(1 - e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon} + \\
&e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon}(|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon - |\sqrt{P}x_{kp}|^2/\epsilon)\Big) \\
&+ \lambda\sqrt{P}\left\|\mathbf{y_p} - \mathbf{H}\mathbf{x_p}\right\|_2^2 .
\end{aligned} \tag{17}$$

Calculate the first order partial derivative of the above equation and we get:

$$\mathbf{x}_p^{IR} = ((\lambda\sqrt{P})^{-1}\mathbf{D_1}^{(t)} + \mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{Y}_p, \tag{18}$$

where $\mathbf{x}_p^{IR}$ is the detection signal at the pth symbols of K users as calculated by the IR algorithm. $\mathbf{Y}_p$ is the pth elements of matrix Y. $\mathbf{D_1}$ is a diagonal matrix:

$$\mathbf{D_1}^{(t)} \cong \begin{bmatrix} e^{|\sqrt{P}\widetilde{x}_{1p}^{(t)}|^2/\epsilon}, & & \\ & \cdots & \\ & & e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon} \end{bmatrix}. \tag{19}$$

B. Log-Sum form

The Log-Sum functions are widely used in the context of dealing with log-likelihood functions and convex optimisation problems. we introduce the Log-Sum objective function using the same approach:

$$\begin{aligned}
&\min_{\boldsymbol{X}} F(\boldsymbol{X}) = \sum_{k=1}^{K} \log(||X_k||^2 + \epsilon) \\
&\text{s.t.} \quad ||\boldsymbol{Y} - \boldsymbol{H}\boldsymbol{X}||_F \leq \eta,
\end{aligned} \tag{20}$$

and the resulting Lagrange equation is expressed as follows:

$$\min_{\mathbf{X}} L(\mathbf{X}) \triangleq \sum_{k=1}^{K} \log\left(\left\|\mathbf{X}_k\right\|_2^2 + \varepsilon\right) + \lambda \left\|\mathbf{Y} - \mathbf{H}\mathbf{X}\right\|_F^2 . \tag{21}$$

The surrogate function is deduced in the same way:

$$\begin{aligned}
&Q(\mathbf{X}|\tilde{\mathbf{X}}^{(t)}) \\
&\triangleq \sum_{k=1}^{K} \left(\frac{\left|\sqrt{P}x_{kp}\right|^2 + \varepsilon}{\left|\sqrt{P}\widetilde{x}_{kp}^{(t)}\right|^2 + \varepsilon} + \log\left(\left|\sqrt{P}\widetilde{x}_{kp}^{(t)}\right|^2 + \varepsilon\right) - 1 \right),
\end{aligned} \tag{22}$$

the optimization problem can be formulated as follows:

$$\begin{aligned}
&\min_{\mathbf{X}} L(\mathbf{X}) \\
&\triangleq \sum_{k=1}^{K} \left(\frac{\left|\sqrt{P}x_{kp}\right|^2 + \varepsilon}{\left|\sqrt{P}\widetilde{x}_{kp}^{(t)}\right|^2 + \varepsilon} + \log\left(\left|\sqrt{P}\widetilde{x}_{kp}^{(t)}\right|^2 + \varepsilon\right) - 1 \right) \\
&+ \lambda \left\|\mathbf{Y} - \mathbf{H}\mathbf{X}\right\|_F^2 .
\end{aligned} \tag{23}$$

Iterative reduction of the above equations using the IR method easily yields smooth iterative results:

$$\mathbf{x}_p^{IR} = ((\lambda\sqrt{P})^{-1}\mathbf{D_2}^{(t)} + \mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{Y}_p, \tag{24}$$

where:

$$\mathbf{D_2}^{(t)} \cong \begin{bmatrix} \frac{1}{\left|\left|\sqrt{P}\tilde{x}_{1p}^{(t)}\right|^2+\varepsilon\right|}, & & \\ & \cdots & \\ & & \frac{1}{\left|\left|\sqrt{P}\tilde{x}_{kp}^{(t)}\right|^2+\varepsilon\right|} \end{bmatrix}. \tag{25}$$

Algorithm 1 The IR-MIMO Algorithm

Input: $x^{(t)}$: the value of the tth iteration; y_p: the observed value; $\mathbf{H}$: the matrix contains channel information; $Iter$: number of iterations; $Iter_{max}$: the max number of iterations;

Output: optimal $x^{(Iter)}$;

Initialize $x^{(0)}$ as an all one column vector; initialize $x^{(1)}$ as an all zero column vector: $\eta = 0.1$; $\varepsilon = 0.1$; $Iter = 0$;
while $Iter_{max} > Iter$ **do**
  $Iter = Iter + 1$;
  for $i = 1 : P$ **do**
    Compute $D_1 = \text{diag}(e^{|\sqrt{P}\widetilde{x}_{kp}^{(t)}|^2/\epsilon})$, or $D_2 = \text{diag}(\frac{1}{\left|\left|\sqrt{P}\tilde{x}_{kp}^{(t)}\right|^2+\varepsilon\right|})$;
    $y_p = \mathbf{Y}_p$;
    $\mathbf{x}_p^{IR-exp} = ((\lambda\sqrt{P})^{-1}\mathbf{D_1}^{(t)} + \mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{Y}_p$ or
    $\mathbf{x}_p^{IR-log} = ((\lambda\sqrt{P})^{-1}\mathbf{D_2}^{(t)} + \mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{Y}_p$;
  end for
  if $\left\|x^{(Iter+1)} - x^{(Iter)}\right\|^2 < 10^{-6}$ **then**
    break;
  end if
end while

The algorithm is organized as **Algorithm 1**. In summary, we solve the non-convex optimisation problem by using two surrogate functions that satisfy the conditions of the MM algorithm by iteratively searching for their lower bounds, and we can ultimately obtain our optimal solution by using (18), (19), (24), (25). It should be noted that the proposed method is well suited for sparse scenarios and **does not require a priori information about the user activity factor**.

IV. Simulation results

In this section, we verify the effectiveness of the proposed iterative reweighed schemes under the uplink Grant-Free MIMO-NOMA system. Depending on the needs of the network and specific application scenarios, this work can be applied to both single and multi-cells. To simplify, we only consider a single-cell MU-MIMO system, which consists of K single-antenna users and a BS with M antennas. In order to capture the sporadic nature of massive user access, we define a variable p to manifest the sparsity of user access.

Consider a Gaussian complex channel and use the symbol error rate (SER) calculated from the recovered signals of multiple transmissions as a criterion for the performance of the algorithm:

$$SER = \frac{\sum_{i=1}^{RPT}\sum_{j=1}^{P}\sum_{z=1}^{K} X_{ijz} \neq \tilde{X}_{ijz}}{K \times \mathrm{RPT}}. \quad (26)$$

where RPT denotes the repeat numbers of each Monte Carlo simulation. X_{ijz} represents the jth column and zth row element of the transmit signal at ith RPT result, while $\tilde{X}_{ijz}$ represents the jth column and zth row element of the recovery signal at ith RPT result.

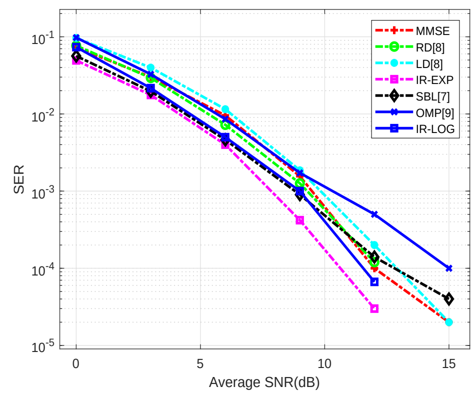
We set parameter $p = 0.2$ to reflect the low activity feature of users, and compare the proposed IR algorithm in exponential form (IR-EXP) and IR algorithm in Log-Sum form (IR-LOG) with Linear minimax mean-square error (MMSE), the Lasso detector (LD), the Ridge detector (RD) , the orthogonal matching pursuit (OMP) detector and the SBL detector under the antennas number $M = 8$, $M = 16$ and the device number $K = 5$, $K = 64$. As Fig.1 shows, the proposed IR algorithm surpasses other MUD algorithms due to the MM algorithm's robust global convergence promotion. By leveraging convex optimization properties and employing surrogate functions, it facilitates the exploration of sub-optimal or nearly optimal solutions. This iterative approach enables weight adjustments based on prior estimates in each iteration, allowing highly-weighted observations to contribute significantly to the estimation process while mitigating the influence of outliers. Consequently, it facilitates convergence to more accurate results.

According to the calculation, the complexity of the algorithms is $\mathcal{O}(K^3)$, with the exception of LD, which has a complexity of polynomial time. To analyze the computational complexity of the algorithm, we recorded the running time for the proposed algorithms and compared with other other algorithms in Table I, it is evident that the IR scheme exhibits comparatively lower complexity, ensuring effectiveness when compared to other traditional MUD algorithms.

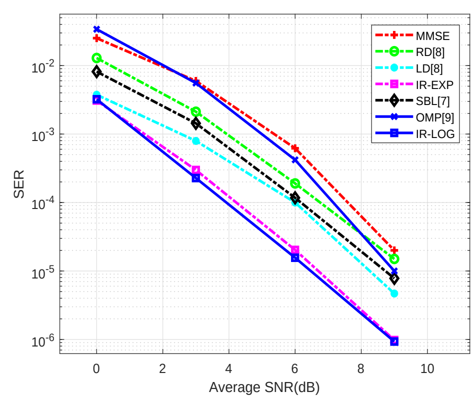
In addition, we have analysed the convergence rate of the proposed algorithm in Fig.2 (a) and compared it with its counterpart SBL which is also an iterative algorithm. It can be observed that as the number of iterations increases, both forms of our proposed IR algorithms converge much faster than the SBL algorithm, and the IR-LOG algorithm converges the fastest. As the regularisation constant increases, the ability to constrain the parameters becomes stronger, making the change in values less drastic and thus achieving rapid convergence. As Fig.2 (b) shows, a comparison was conducted to assess the recovery performance of the algorithms in the three cases of regularisation constants. The results demonstrated that the EXP form of IR performed optimally with large regularisation constants, whereas the LOG form exhibited the opposite trend. All algorithms exhibited convergence in one or two iterations, indicating that the algorithm's ability to effectively reach the hard judgement threshold in the first iteration.

TABLE I: Running time of MUD algorithms under $M = 16$ and $K = 64$ (second).

Algorithm	runtime($p_a = 0.1$)	runtime($p_a = 0.2$)
MMSE	0.001422	0.001068
RD	0.001048	0.000933
LD	1.001371	1.045752
OMP	0.710469	0.849271
SBL	**21.527146**	**21.562025**
IR-LOG	**0.013215**	**0.016604**
IR-EXP	**0.010205**	**0.011349**



(a)

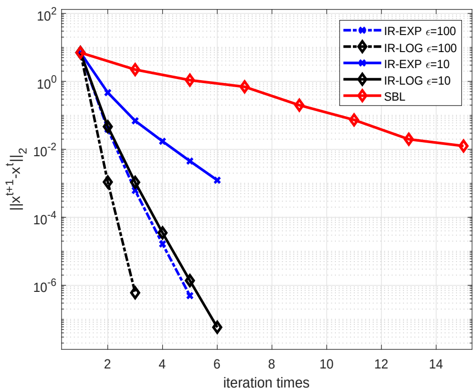


(b)

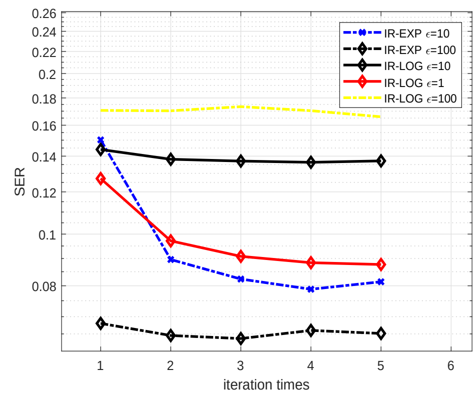
Fig. 1: SER performance of different algorithms at (a) $M = 8$ and $K = 5$; (b) $M = 16$ and $K = 64$.

V. Conclusions

In this paper, we consider the use of an IR scheme combined with the MM algorithm to recovery sparse signals under the Grant-Free MIMO-NOMA model without the need for a priori information, and two forms of the IR scheme are considered. Simulation results demonstrate the superior performance of the proposed IR scheme. Our multi-user detection algorithm has a wide range of potential applications, and in the future research, it can be applied to Cell-free [16], [17] networks to further improve system performance and user experience.



(a)



(b)

Fig. 2: (a) The convergence rate of the proposed algorithm; (b) Algorithm performance with number of iterations at $SNR = 0$, $M = 8, N = 5$.

References

[1] Z. Lin, M. Lin, T. de Cola, J. -B. Wang, W. -P. Zhu and J. Cheng, ”Supporting IoT With Rate-Splitting Multiple Access in Satellite and Aerial-Integrated Networks,” in *IEEE Internet of Things Journal*, vol. 8, no. 14, pp. 11123-11134, July, 2021.

[2] Z. Lin, M. Lin, B. Champagne, W. -P. Zhu and N. Al-Dhahir, “Secrecy-Energy Efficient Hybrid Beamforming for Satellite-Terrestrial Integrated Networks,” in *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 6345-6360, Sept. 2021.

[3] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, F. Tufvesson, “Scaling up MIMO: opportunities and challenges with very large arrays,” in *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40-60, Jan. 2013.

[4] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin and Z. Wang, “Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends,” in *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74-81, Sept. 2015.

[5] T. L. Marzetta, “Noncooperative cellular wireless with unlimited numbers of base station antennas,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[6] S. Bhat and A. Chockalingam, “Sparsity-exploiting detection of large-scale multiuser GSM-MIMO signals using FOCUSS,” in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, Nanjing, China, 2016.

[7] H. Iimori, T. Takahashi, K. Ishibashi, G. T. F. de Abreu and W. Yu, “Grant-Free access via bilinear inference for cell-free MIMO with low-coherence pilots,” in *IEEE Transactions on Wireless Communications*, vol. 20, no. 11, pp. 7694-7710, Nov. 2021.

[8] F. Li, Y. Peng and W. Chen, “Sparse Multi-User Detection with imperfect channel estimation for NOMA systems under non-Gaussian noise,” in *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 246-250, Feb. 2021.

[9] X. Zhang, P. Fan, J. Liu, and L. Hao, “Bayesian learning-based multiuser detection for grant-free NOMA systems,” *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 6317-6328, Aug. 2022.

[10] H. Zhu and G. B. Giannakis, “Exploiting sparse user activity in multiuser detection, *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 454-465, Feb. 2011.

[11] S. Adnan, Y. Fu, N. Junejo, Z.Chen, and H. Esmaiel, “Sparse detection with orthogonal matching pursuit in multiuser uplink quadrature spatial modulation MIMO system”, *IET Communications*, vol. 13, no. 20, pp. 3472-3478, 2019.

[12] X. Zhang, F. Labeau, Y. -C. Liang and J. Fang, “Compressive Sensing-based Multiuser Detection via iterative reweighed approach in M2M communications,” in *IEEE Wireless Communications Letters*, vol. 7, no. 5, pp. 764-767, Oct. 2018.

[13] X. Zhang, L. Hao, P. Fan, J. Liu and L. Yang, “Iterative reweighed approach for Multiuser Detection with multiple measurement vector in MTC communications,” *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, Antwerp, Belgium, 2020.

[14] E. Cands, M. Wakin, and S. Boyd, “Enhancing sparsity by reweighted l1 minimization,” *J. Fourier Anal. Appl.*, vol. 14, no. 5, pp. 877-905, Dec. 2008.

[15] M. Figueiredo, “Adaptive sparseness for supervised learning,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1150-1159, Sept. 2003.

[16] X. Pan, Z. Zheng, Z. Fei and Z. Zhang, ”On the Design of Optimal One-Shot Distributed Combining for Cooperative Multi-UAV Systems,” in IEEE Transactions on Vehicular Technology, vol. 72, no. 5, pp. 6827-6832, May 2023.

[17] E. Bjrnson and L. Sanguinetti, ”Making Cell-Free Massive MIMO Competitive With MMSE Processing and Centralized Implementation,” in IEEE Transactions on Wireless Communications, vol. 19, no. 1, pp. 77-90, Jan. 2020.