User Grouping with Fairness Consideration in Massive MIMO System with Varying Channel Quality

Shengping Zhou, Ming Zeng, Zesong Fei
School of Information and Electronics, Beijing Institude of Technology
Beijing, China
mzengzm@bit.edu.cn

***Abstract*—In recent years, the demand for higher communication capacity in wireless networks has led to the development of massive multiple-input multiple-output (MIMO) systems. Although there have been some researches on addressing the multi-user resource allocation problem in these systems, they are not suitable for the situation where the channel quality of users varies greatly. To solve the resource allocation problem with fairness consideration in the context of high variation in user channel qualities, we propose a new resource allocation method based on user grouping. We first calculate the average channel quality (ACQ) of each user and group them accordingly. Then, different allocation methods are adopted for users with varying channel qualities. Additionally, we found that using the sum of logarithms of every user's rate (SLR) as the fairness metric can adapt to different situations automatically. Simulation results demonstrate that the proposed algorithm achieves high system throughput while maintaining fairness.**

***Index Terms*—SLR, massive MIMO, resource allocation, user grouping**

I. Introduction

In recent years, the demand for wireless data transmission capacity has been growing exponentially. To address this problem, a sophisticated technology called multi-user multiple-input multiple-output (MU-MIMO) has been proposed [1]. MU-MIMO systems enable users to concurrently reuse the same time-frequency resources at the base stations, significantly improving the throughput of communication systems. However, the performance of MU-MIMO systems is limited by the number of antennas at the base stations (BSs) [2], which has hampered the development of massive MIMO systems. The performance of massive MIMO systems can be improved by simply applying more antennas to the BSs, providing an alternative to traditional methods to increase data transmission capacity such as cell-size shrinking [3]. As a result, massive MIMO systems have been extensively studied to improve spectrum and energy efficiency in wireless systems [4], [5].

To further improve system throughput, researches have been conducted on resource allocation algorithms for massive MIMO systems. A low-complexity algorithm to maximize sum-rate was proposed in [6], while the user experience is not considered. So in [7], users are divided into two groups (or clusters), i.e., quality of service (QoS) users and non-QoS users, and two successive convex approximation algorithms are used to guarantee the QoS for QoS users and maximize the sum-rate of non-QoS users, simultaneously. But this may result in a poor experience for non-QoS users.

Therefore, fairness was taken into account in resource allocation in [8] and [9]. In [8], a genetic based user grouping algorithm is proposed with the aim of ensuring fairness through an equitable allocation of communication resources to each user. However, in scenarios where substantial variations in channel quality exist among users, the allocation of identical communication resources may still result in considerable disparities in communication rates among users. In [9] and [10], proportional rate constraints are utilized to guarantee fairness. But this will result in a computationally complex NP-hard problem in their algorithms when there are significant differences in channel quality among users, because setting target rates for proportional rate constraints is a non-convex linear programming problem, and the solution to this issue is not delineated in [9] and [10].

To solve the resource allocation problem in the context of high variation in user channel qualities, in this letter, we propose a joint user grouping and resource allocation algorithm with fairness evaluated by SLR for downlink time division duplex (TDD) massive MIMO systems. Taking the difference in channel quality among users and the spatial correlation between users into account, the proposed algorithm can optimize system throughput and SLR under an acceptable level of computational complexity. Specifically, our algorithm first evaluates each user's channel quality based on the downlink channel matrix, and then groups users based on their channel quality. We then propose two resource allocation algorithms for different user groups, with fairness constraints to allocate resource block (RB) to proper users. Simulation results show that our algorithm achieves good performance in terms of both system throughput and fairness.

The main contributions of this paper are as follows:

1) By grouping users and adopting different allocation methods for different groups, the proposed algorithm can

effectively handle the issue of varying channel qualities among users. To the best of our knowledge, this is the first work to address the user allocation problem with fairness consideration under the premise of vastly differing channel qualities among users.

2) We introduce SLR as a replacement for proportional rate constraints as a fairness metric, which solves the problem of setting target rates for users in complex environments.

The remainder of this paper is organized as follows: Section II outlines the system model and problem formulation in the massive MIMO system. Section III proposes our resource allocation algorithm based on user grouping. In Section IV, we give the simulation results of the proposed algorithm, and verify the performance of our proposed algorithm compared with existing algorithms. Finally, we summarize the paper in the concluding section.

II. System Model and Problem Formulation

As shown in Fig.1, we consider a single-cell massive MIMO communication system consisting of K users and N RBs per time transmission interval (TTI), where the base station is equipped with N_T antennas. In each TTI, multiple users can transmit information on the same RB through space division multiplexing.

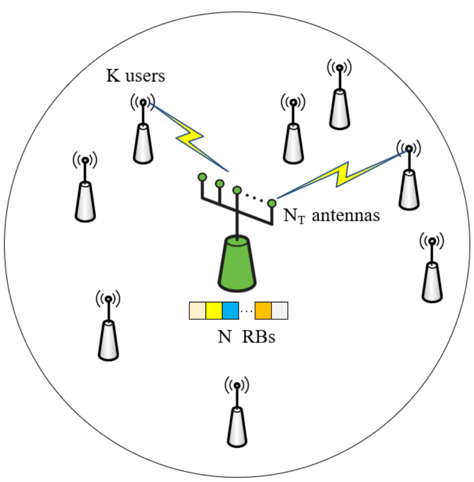

Fig. 1. A single-cell massive MMIO communication system consisting of a N_T-antennas BS and K single-antenna users with N RB resources.

A. System Model and Zero-Forcing Method

The channels of all users on RB n can be represented as a channel matrix

$$\mathbf{H}_n = [\mathbf{h}_{1,n}, \cdots, \mathbf{h}_{k,n}, \cdots, \mathbf{h}_{K,n}]^{\mathrm{H}}, \tag{1}$$

where $\mathbf{h}_{k,n} \in \mathbb{C}^{N_T \times 1}$ is the channel of user k on RB n. The base station performs weighted transmission on RB n through the precoding matrix can be expressed as

$$\mathbf{W}_n = [\mathbf{w}_{1,n}, \cdots, \mathbf{w}_{k,n}, \cdots, \mathbf{w}_{K,n}], \tag{2}$$

which can be given by the zero-forcing method [11] as

$$\mathbf{W}_n = \mathbf{H}_n^{\mathrm{H}}(\mathbf{H}_n\mathbf{H}_n^{\mathrm{H}})^{-1}. \tag{3}$$

And the vector $\mathbf{w}_{k,n} \in \mathbb{C}^{N_T \times 1}$ represents the precoding vector for user k on RB n. To normalize it for subsequent calculations, we can use the equation as follows:

$$\widetilde{\mathbf{w}}_{k,n} = \frac{\mathbf{w}_{k,n}}{\|\mathbf{w}_{k,n}\|}. \tag{4}$$

According to Shannon's formula, the max data transmission rate $R_{k,n}$ for user k on RB n can be expressed by

$$R_{k,n} = B\log_2(1 + SINR_{k,n}), \tag{5}$$

where B represents the channel bandwidth. As a result, increasing $R_{k,n}$ is equivalent to increasing $SINR_{k,n}$. $SINR_{k,n}$ is the Signal-to-Interference-plus-Noise-Ratio for user k on RB n, which can be given by

$$SINR_{k,n} = \frac{\left|\mathbf{h}_{k,n}^H \widetilde{\mathbf{w}}_{k,n} p_{k,n}\right|^2}{\sum_{l \in K_n \backslash \{k\}} \left|\mathbf{h}_{k,n}^H \widetilde{\mathbf{w}}_{l,n} p_{l,n}\right|^2 + B\sigma^2}, \tag{6}$$

where σ^2 represents the power of the white Gaussian noise, $p_{k,n} = \sqrt{P_{k,n}}$ represents the voltage value of the transmitted signal for user k on RB n, where $P_{k,n}$ represents the power of the transmitted signal.

As the zero-forcing method can effectively eliminate inter-user interference, we can mathematically express it as $\mathbf{h}_{k,n}^H \widetilde{\mathbf{w}}_{l,n} = 0, l \neq k$. Thus, we can approximate formulation (6) as follows:

$$\begin{aligned} SINR_{k,n} &= \frac{\left|\mathbf{h}_{k,n}^H \widetilde{\mathbf{w}}_{k,n} p_{k,n}\right|^2}{\sum_{l \in K_n \backslash \{k\}} \left|\mathbf{h}_{k,n}^H \widetilde{\mathbf{w}}_{l,n} p_{l,n}\right|^2 + B\sigma^2} \\ &\approx \frac{\left|\mathbf{h}_{k,n}^H \widetilde{\mathbf{w}}_{k,n} p_{k,n}\right|^2}{B\sigma^2} \end{aligned} \tag{7}$$

B. Evaluation Index and Problem Formulation

In addition to throughput, fairness is also an important metric in resource allocation algorithms. Existing researches usually use proportional rate constraints [10] to ensure fairness

$$F_P = \frac{\left(\sum_{k=1}^{K} R_k/\lambda_k\right)^2}{K \sum_{k=1}^{K} (R_k/\lambda_k)^2}, \tag{8}$$

where R_k is the user's average data rate during T TTIs and λ_k is the target data rate for user k. F_P can be used to indicate the proximity of throughput proportions among users to the preset target rate $\{\lambda_k\}_{k=1}^K$. However, in scenarios where users' channel qualities exhibit significant variations, it becomes necessary to allocate reasonable $\{\lambda_k\}_{k=1}^K$ for each user, which is a computationally complex NP-hard problem. Therefore, we aim to address the challenge of setting the target rate ratio for different users by using another evaluation index called SLR, which can be represented as follows:

$$\mathrm{SLR} = \sum_{k=1}^{K} \log_2\left(\sum_{n=1}^{N} x_{k,n} R_{k,n}\right), \tag{9}$$

where $x_{k,n}$ denote a binary indicator variable indicating whether user k transmits information on RB n. Mathematically, this can be expressed as follows:

$$x_{k,n} = \begin{cases} 1, & \text{if user } k \text{ is on } K_n \\ 0, & \text{if user } k \text{ is not on } K_n \end{cases}, \tag{10}$$

where K_n denotes the set of users who transmit information on RB n. Using SLR as a fairness metric is obviously reasonable because SLR can only achieve its maximum value when each user receives corresponding communication resources in a way that is fair relative to their channel conditions. The more unfair the allocation, the lower the value of SLR. And by using SLR, the problem of setting $\{\lambda_k\}_{k=1}^{K}$ in proportional fairness is avoided, making the proposed algorithm suitable for complex scenarios where there are significant differences in channel quality among users.

With the goal of balancing the trade-off between system throughput and fairness, we can formulate the problem as follows:

$$\max_{x_{k,n},P_{k,n}} ((\sum_{k=1}^{K}\sum_{n=1}^{N} x_{k,n}R_{k,n}) + I\times \sum_{k=1}^{K} \log_2(\sum_{n=1}^{N} x_{k,n}R_{k,n}))$$
$$\text{s.t.} I \geq 0$$
$$\sum_{k=1}^{K}\sum_{n=1}^{N} P_{k,n} \leq P. \tag{11}$$

where I represents the importance of SLR in relation to system throughput, and P represents the maximum transmit power of the entire base station. To simplify the problem, we assume that P is equally allocated to all RBs, which can be expressed as

$$\sum_{k=1}^{K} P_{k,n} \leq \frac{P}{N}, \quad n = 1,2,3,\cdots,N \tag{12}$$

And for each RB, we adopt uniform power allocation among different users, which is written as:

$$P_{i,n} = P_{j,n}, \quad i,j = 1,2,3,\cdots,K, \quad n = 1,2,3,\cdots,N. \tag{13}$$

Under this simplification, the problem (11), which becomes an NP-hard combinatorial optimization problem [12], is still difficult to solve, and the only way to obtain the optimal solution is through exhaustive search. However, the complexity of exhaustive search reaches 2^{KN}, which is completely unacceptable in practical applications. In the following section, we propose a suboptimal resource allocation algorithm.

III. The Proposed Resource Allocation Algorithm

According to (7), we can observe that the value of $SINR_{k,n}$ mainly depends on B, σ^2, $h_{k,n}^{\mathrm{H}}$, $\mathbf{w}_{k,n}$ and $p_{k,n}$. Among these parameters, B, σ^2 and $h_{k,n}^{\mathrm{H}}$ are fixed values determined by the communication environment. It can be seen from (4) that the value of $\widetilde{\mathbf{w}}_{k,n}$ is directly determined by $\mathbf{w}_{k,n}$, which is the constituent element of $\mathbf{W}_n$. Furthermore, considering (3), we can see that the value of $\widetilde{\mathbf{w}}_{k,n}$ is also determined by $\mathbf{H}_n$, which is also a fixed value determined by the communication environment. Therefore, the $SINR_{k,n}$ can be only improved by optimizing $p_{k,n}$. By substituting $p_{k,n} = \sqrt{P_{k,n}}$ and (7) into Shannon's formulation, we can obtain:

$$R_{k,n} = B\log_2(1 + \frac{\left|\mathbf{h}_{k,n}^{H}\widetilde{\mathbf{w}}_{k,n}\right|^2}{B\sigma^2}P_{k,n}), \tag{14}$$

take the partial derivative of (14) into account, the value of $\frac{\partial R_{k,n}}{\partial P_{k,n}}$ can be expressed as

$$\frac{\partial R_{k,n}}{\partial P_{k,n}} = \frac{\left|\mathbf{h}_{k,n}^{H}\widetilde{\mathbf{w}}_{k,n}\right|^2}{(\ln 2)\sigma^2}\frac{1}{1+\frac{\left|\mathbf{h}_{k,n}^{H}\widetilde{\mathbf{w}}_{k,n}\right|^2}{B\sigma^2}P_{k,n}}, \tag{15}$$

formula (15) can be approximated as

$$\frac{\partial R_{k,n}}{\partial P_{k,n}} = \begin{cases} \frac{\partial R_{k,n}}{\partial P_{k,n}}_{const} = \frac{\left|\mathbf{h}_{k,n}^{H}\widetilde{\mathbf{w}}_{k,n}\right|^2}{(\ln 2)\sigma^2}, & \frac{\left|\mathbf{h}_{k,n}^{H}\widetilde{\mathbf{w}}_{k,n}\right|^2}{B\sigma^2}P_{k,n} \ll 1, \\ \frac{\frac{\partial R_{k,n}}{\partial P_{k,n}}_{const}}{\frac{\left|\mathbf{h}_{k,n}^{H}\widetilde{\mathbf{w}}_{k,n}\right|^2}{B\sigma^2}P_{k,n}}, & \frac{\left|\mathbf{h}_{k,n}^{H}\widetilde{\mathbf{w}}_{k,n}\right|^2}{B\sigma^2}P_{k,n} \gg 1, \\ \frac{\frac{\partial R_{k,n}}{\partial P_{k,n}}_{const}}{1+\frac{\left|\mathbf{h}_{k,n}^{H}\widetilde{\mathbf{w}}_{k,n}\right|^2}{B\sigma^2}P_{k,n}}, & else. \end{cases} \tag{16}$$

The above formula suggests that the energy should be allocated preferentially to transmission channels with better conditions to obtain a higher $R_{k,n}$ when the value of $\frac{\left|\mathbf{h}_{k,n}^{H}\widetilde{\mathbf{w}}_{k,n}\right|^2}{B\sigma^2}P_{k,n}$ is much less than 1. As the energy allocated to the channels with better transmission conditions, the value of $\frac{\partial R_{k,n}}{\partial P_{k,n}}$ gradually decreases. When this value is reduced below a preset threshold, the energy should be properly allocated to other channels with relatively poor transmission conditions.Hence, when the channel quality of a user is poor, the energy allocation should be concentrated in the channel with the best transmission conditions as much as possible, while for better channel conditions, the energy should be distributed among multiple channels.

In addition, we use average channel quality (ACQ) to evaluate the channel quality of each user, which can be expressed as:

$$ACQ_k = \frac{\sum_{n=1}^{N}\sqrt{\mathbf{h}_{k,n}^{H}\mathbf{h}_{k,n}}}{N} \tag{17}$$

In this proposed algorithm, the energy allocated to each user on a RB is indirectly controlled by adjusting the number of users allocated to the RB. For users with poor channel quality, it is required to occupy more energy on RBs with better transmission conditions to improve data transmission efficiency. Therefore, RBs carrying users with poor channel quality should be assigned with fewer users compared to other RBs.

However, if users with better channel quality are allocated to RBs carrying fewer users, the data transmission efficiency will decrease due to the decrease in $\frac{\partial R_{k,n}}{\partial P_{k,n}}$. Consequently, users with poor channel quality are grouped together, rather than mixed with users with better channel quality, and RBs with the best transmission conditions relative to the users in the group are selected for these users earlier because changes in channel quality have a greater impact on these users.

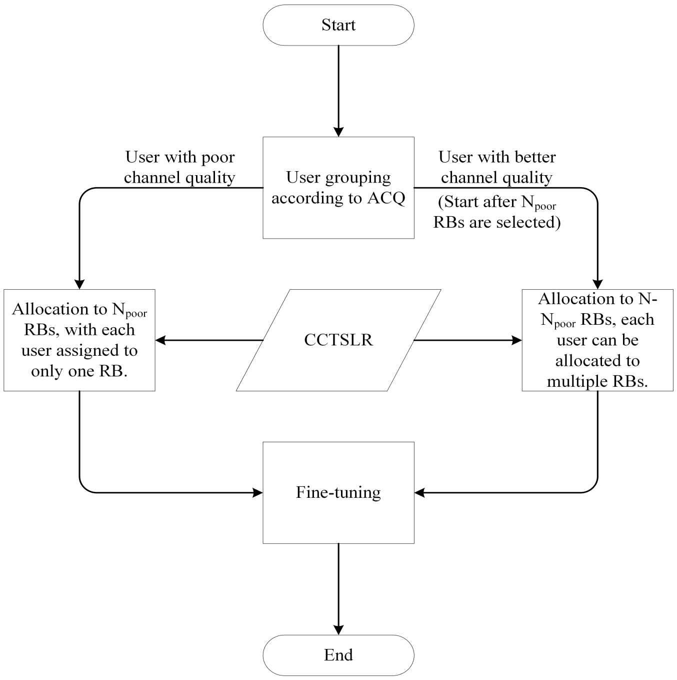

Fig. 2. Proposed resource allocation algorithm

As shown in Fig.2, based on the above analysis, our proposed algorithm can be implemented through the following steps:

Step 1: Obtain maximum transmission power P, the power of white Gaussian noise channel σ^2, matrix $\mathbf{H}$, user set $\mathbb{U}$ and RB set $\mathbb{R}$. It is assumed that $\mathbf{H}$ can be obtained through the feedback from the previous TTI.

Step 2: The ACQ of each user is calculated according to $\mathbf{H}$ and formula (17), and the users are grouped into those with better channel quality and those with poor channel quality according to the value of ACQ.

Step 3: Allocate resources to the users in the poor channel quality group. These users will be allocated to N_{poor} RBs, with each user is assigned with only one RB.

Step 4: Allocate the remaining $N - N_{poor}$ RBs to the users with better channel quality, and each user can be allocated to multiple RBs.

Step 5: Fine-tune. Repeat F times : Remove L allocations between users and RBs with the lowest performance gains, as well as randomly deleting D allocations. After the removal, reallocate users to RBs.

Steps one to four decompose the problem into multiple sub-problems and solve them gradually, considering only the optimal solution for the current sub-problem. Therefore, some allocations may appear unreasonable from a global perspective. Hence, in step five, we perform fine-tuning to remove these allocations and correct such errors. Moreover, we randomly delete some allocations to aid the algorithm in exploring more possibilities in the solution space and escape from local optimal solutions. During the implementation of the algorithm, the performance evaluation criterion is given by (11), which we refer to as comprehensive considerations of throughput and sum of the logarithm of every user's Rate ($CCTSLR$) hereinafter.

The detailed implementation process of the entire algorithm is presented in Algorithm 1.

Algorithm 1 Resource allocation algorithm based on user grouping

Input: maximum transmission power P, the number of RBs for users with poor channel quality N_{poor}, the number of repetitions times F, the number of deletions L, the number of random deletions D, channel matrix $\mathbf{H}$, user set $\mathbb{U}$ and RB set $\mathbb{R}$

Output: the user set for each RB $\mathbb{S}_1$, $\mathbb{S}_2$, $\mathbb{S}_3$,......, $\mathbb{S}_N$(The set of these user sets is $\mathbb{S}_{total}$), the value of $CCTSLR$

initialize $\mathbb{S}_1 = \mathbb{S}_2 = \mathbb{S}_3 ...... = \mathbb{S}_N = \emptyset$.
for each user, calculate ACQ. Divide $\mathbb{U}$ into $\mathbb{U}_p$(with poor channel quality) and $\mathbb{U}_b$(with better channel quality) based on ACQ
for i = 1 to N_{poor} **do**
  find user $k^* = \arg\max_{k \in \mathbb{U}_p}(ACQ_k)$
  for each RB in $\mathbb{R}$ **do**
    $\mathbb{S}_n = k^*$
  end for
  $\mathbb{U}_p = \mathbb{U}_p \backslash \{k^*\}$
  for each RB n in $\mathbb{R}$ **do**
    set $\mathbb{U}_p^{temp} = \mathbb{U}_p$, $\mathbb{S}_n^{temp} = \emptyset$
    for j = 2 to C_i **do** ▷ C_i is the number of users on the RB selected in the i-th cycle
      find user $k_{temp} = \arg\max_{k \in \mathbb{U}_p^{temp}}(CCTSLR)$, $\mathbb{S}_n^{temp} = \mathbb{S}_n^{temp} \cup \{k_{temp}\}$
      $\mathbb{U}_p^{temp} = \mathbb{U}_p^{temp} \backslash \{k_{temp}\}$
    end for
  end for
  find user set $\mathbb{S}_n^{temp} = \arg\max_{\mathbb{S}_n^{temp} \in \mathbb{S}_{total}^{temp}} (CCTSLR)$, $\mathbb{S}_n = \mathbb{S}_n^{temp}$
  $\mathbb{R} = \mathbb{R} \backslash \{n\}$
end for

for each user k in $\mathbb{U}_b$ **do**
  find $n_{temp} = \arg\max_{n \in \mathbb{R}}(CCTSLR)$, $\mathbb{S}_{n_{temp}} = \mathbb{S}_{n_{temp}} \cup k$
  ▷ Prevent users from not being assigned to RB
end for

for each RB n in $\mathbb{R}$ **do**

$\mathbb{U}_b^n = \mathbb{U}_b$
end for

for $\mathbb{R} \neq \emptyset$ **do**
  for each RB n in $\mathbb{R}$ **do**
    find $k_{temp} = \arg\max_{k \in \mathbb{U}_b^n} (CCTSLR)$
    if the value of $CCTSLR$ rises after assigning user k_{temp} to RB n **then**
      $\mathbb{S}_n = \mathbb{S}_n \cup k_{temp}$, $\mathbb{U}_b^n = \mathbb{U}_b^n \backslash \{k_{temp}\}$
    else
      $\mathbb{R} = \mathbb{R} \backslash \{n\}$
    end if
  end for
end for

for f = 1 to F **do**
  for l = 1 to L **do**
    remove the allocation between users and RBs with the lowest performance gains.
  end for
  for d = 1 to D **do**
    randomly deleting an allocation.
  end for
  reallocate users to RBs
end for

IV. Simulation Results and Analysis

The proposed algorithm is compared with the existing resource allocation algorithms. It should be noted that the simulation parameters N_{poor}, F, L, and D are not fixed, which are changed according to user channel conditions, computing resource, capacity requirements etc..

The dataset used in our experiments is the Massive_MIMO_Dataset from https://www.huaweicloud.com. The main feature of this dataset is the significant differences in channel qualities among users. The results in the figures are averaged under ten different channel conditions, each averaged over one hundred TTIs to provide a reliable simulation performance. In this paper, we set N=8, N_T=32, and assume B is invariable. Table I shows an example of the allocation between users and RBs.

TABLE I
An Example of the Allocation Between Users and RBs

RB \ User	1	2	3	...	K
1	1	0	1	...	1
2	1	0	1	...	0
3	0	0	1	...	0
4	1	0	0	...	1
5	0	1	0	...	1
...	...	...	...	$x_{k,n}$	...
N	1	1	0	...	1

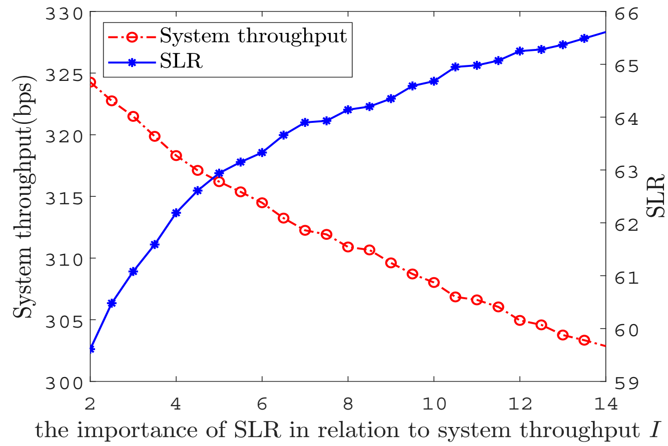

Fig. 3. System throughput and SLR vs I

Fig.3 illustrates the performance of the proposed algorithm for K=24 under different values of I. The results indicate that as the value of I increases, the SLR, which reflects both fairness and user experience, also increase, However, the system throughput decrease. Accordingly, we can achieve an optimal trade-off between the overall system throughput and the user fairness at $I = 5$.

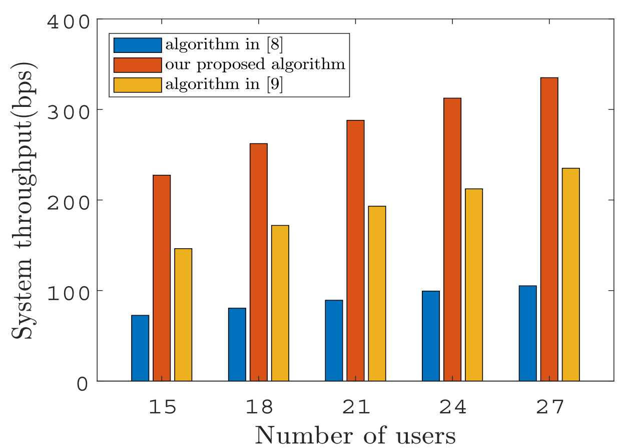

Fig. 4. System throughput vs number of users

Fig.4 presents a comparison of the three algorithms' system throughput growth as the number of users increases when $I = 5$. It is observed that our proposed algorithm has the best performance in terms of system throughput under the environment where there are significant differences in channel qualities among users. This advantage becomes more pronounced with an increasing number of users, since the algorithms proposed in [8] and [9] did not handle the differences in channel quality between users effectively, and allocated each user to RB in the same way, resulting in a decrease in data transmission efficiency and low total system throughput. In contrast, our proposed algorithm groups users based on their channel quality to avoid mixing users with better and poor channel quality. On this basis, the suitable resources allocation algorithms are separately used in these two types of users, which guarantees a superior system throughput.

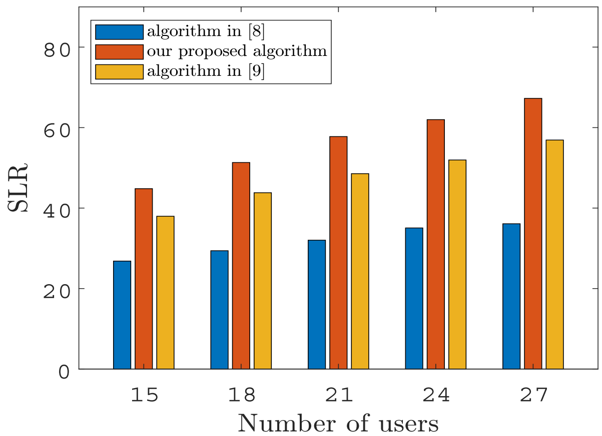

Fig. 5. SLR vs number of users

In Fig.5, it is evident that our proposed algorithm also shows considerable advantages in terms of SLR. This is because the algorithm proposed in [9] requires to set a target rate for each user, which is difficult to achieve due to the varying channel conditions among users in practical scenarios for it's a non-convex linear programming problem. Our algorithm aims at improving CCTSLR and can adapt to different situations automatically without setting a target rate. Additionally, although the algorithm proposed in [8] equitable allocate communication resources to each user, but users with poor channel conditions will get a lower rate throughput, since it requires all users to concentrate energy on one RB. However, our proposed fairness metrics SLR can effectively overcome the shortcomings of above two algorithms, and achieves a great fairness performance under the constraint of the system throughput.

V. Conclusion

To address the resource allocation problem in massive MIMO systems with significant differences in channel qualities among users, we propose a novel algorithm based on user grouping and SLR. Our approach separates users based on their channel quality and allocates resources to them in a differentiated manner to avoid reducing the system throughput due to intermixing of users with varying channel quality. The use of CCTSLR as the sole criterion to determine user-RB allocation leads to superior performance in terms of system throughput, fairness, and user experience. Simulation results confirm the excellent performance of our proposed algorithm. For further research, joint power allocation under multiple TTIs can be considered.

VI. Acknowledgments

This work is supported by National Natural Science Funds of China (Grant No.62001028).

References

[1] C. Lim, T. Yoo, B. Clerckx, B. Lee and B. Shim, "Recent trend of multiuser MIMO in LTE-advanced," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 127-135, Mar. 2013.

[2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742-758, May 2014.

[3] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan and M. C. Reed, "Femtocells: Past, Present, and Future," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497-508, Apr. 2012.

[4] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta and F. Tufvesson., "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013.

[5] L. Sanguinetti, E. Bjrnson and J. Hoydis, "Toward Massive MIMO 2.0: Understanding Spatial Correlation, Interference Suppression and Pilot Contamination," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 232-257, Jan. 2020.

[6] E. Castaneda, A. Silva, R. Samano-Robles and A. Gameiro, "Low complexity user selection for rate maximization in MIMO broadcast channels with downlink beamforming," *Sci. World J*, vol. 2014, no. 2,pp. 983930, Jan. 2014.

[7] H. T. Dao and S. Kim, "Power Allocation for Multiple User-Type Massive MIMO Systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 10965-10974, Oct. 2020.

[8] C. Li, H. Zhu, J. Tang, J. Hu and G. Li, "User Grouping in Multiuser Satellite MIMO Downlink With Fairness Consideration," *IEEE Wireless Commun. Lett.*, vol. 11, no. 8, pp. 1575-1579, Aug. 2022.

[9] X. Yang, S. Zhang, B. Gao and J. Cao, "A Low Complexity Joint User Grouping and Resource Allocation Algorithm in Massive MIMO Systems," *Proc. IEEE Int. Conf. Commun. Technol(ICCT)*, 2019, pp. 914-919.

[10] Zukang Shen, J. G. Andrews and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2726-2737, Nov. 2005.

[11] Q. H. Spencer, A. L. Swindlehurst and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461-471, Feb. 2004.

[12] Y. Sun, D. W. K. Ng, Z. Ding and R. Schober, "Optimal Joint Power and Subcarrier Allocation for Full-Duplex Multicarrier Non-Orthogonal Multiple Access Systems," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1077-1091, Mar. 2017.