A Study on User-antenna Cluster Formation for Cluster-wise MU-MIMO

Sijie Xia†1 Chang Ge†2 Qiang Chen†3 Fumiyuki Adachi†4

†1, 2, 3 *Department of Communications Engineering*, *Tohoku University*
6-6-05 Aza-Aoba, Aramaki, Sendai, Miyagi, 980-8579 Japan

†4 *Research Organization of Electrical Communication*, *Tohoku University*
2-1-1 Katahira, Aoba-ku, Sendai, Miyagi, 980-8577 Japan

E-mail: †1,3,4 {xia-s, chenq, adachi}@ecei.tohoku.ac.jp, †2 ge.chang.q2@dc.tohoku.ac.jp

***Abstract*—The ultra-dense radio access network (RAN) consisting of a large number of distributed antennas and user terminals is considered to improve link capacity and coverage while saving the energy consumption for 5G advanced systems. Because of such a large-scale multi-user multiple-input multiple-output (MU-MIMO), the computational complexity of signal processing is difficult to realize. Aiming to reduce the computational complexity to an acceptable level, the clustering approach is promising to decompose the large-scale MU-MIMO into several smaller-scale ones. K-means clustering is introduced to form either user clusters or antenna clusters and then, cluster member assignment (CMA) is done to perform cluster-wise zero-forcing (ZF) based MU-MIMO. In this paper, we compare through computer simulation possible combinations of two clustering methods and two CMA methods in terms of achievable link capacity.**

Keywords—5G advanced systems, distributed MU-MIMO, Clustering, ZF

I. INTRODUCTION

With the popularization of various kinds of terminal equipment and the growth of mobile users, the aggregate mobile data traffic is increasing yearly. To accommodate such a growing mobile data traffic, the radio access network (RAN) must be densified. Simultaneously, higher frequency bands are an inevitable choice to expand the bandwidth in order to provide reliable service for such a large number of users. Compared with centralized antenna system, distributed antenna system can avoid the problem of the high-frequency electromagnetic wave being often blocked by obstacles through antenna selection while maintaining the base station (BS) coverage area (called cell in this paper). An introduction of a RAN consisting of a large number of distributed antennas (called distributed RAN) is probably an effective approach to improve the energy efficiency and the spectral efficiency simultaneously in 5G advanced systems [1-3].

No matter the antenna deployment is distributed or centralized, for the ultra-dense multi-user multiple-input multiple-output (MU-MIMO), the large-scale channel matrix requires a high level of computational complexity at the BS side. Aiming to reduce the computational complexity, we introduced clustering method to transform the original large-scale MU-MIMO into several smaller-scale ones.

A. Related Works

Clustering or grouping methods for MU-MIMO systems are introduced to improve system performance as in [9-12]. In particular, in [9], authors adopt the thought of clustering in unsupervised learning and provide design schemes for hybrid precoding. Hierarchical-agglomerative-clustering-based scheme to explore the relevance among radio frequency (RF) chains and modified-K-means-based scheme to explore the relevance among antennas in BS are proposed to combat the path loss of millimeter-wave signals in massive MIMO system. Similarly, in order to achieve the massive MIMO gain, [10] proposed a signal-to-interference plus noise ratio (SINR)-based user scheduling and user grouping to obtain a higher sum rate. Moreover, to satisfy the required data rates of future wireless applications, [11] considered a heterogenous network consisting a macro BS and a large number of distributed small cells which can improve local capacity. And distance-based K-means clustering and affinity propagation clustering are proposed to overcome the inter-cell interference which is a dominant limiting factor of system performance. And in [12], authors considered the optimal antenna cluster size in a cell-free frequency division duplex (FDD) large-scale distributed antenna system (DAS) to optimize the system throughput. Here, the antenna cluster-size is determined by locations of users and antennas with a reserved number of users.

B. Contribution

To the best of our knowledge, clustering in dense distributed MU-MIMO with irregular antenna distribution was not considered in the previous studies. Therefore, we try to tackle with this problem to propose a new method. In this paper, we focus on a single-cell area within numerous distributed antennas and users. Aiming to reduce the computational complexity of large-scale MU-MIMO, because each user and antenna are distributed in different positions, we propose a user-antenna clustering (UAC), which consists of the following methods:

- Initial clustering (IC) to form either user clusters or antenna clusters,
- Cluster member assignment (CMA) to allocated users or antennas into the initial clusters.

We apply K-means clustering [4,5] based on the locations of users and antennas to IC. CMA is the significant part of cluster-wise MU-MIMO formation, which also effects the system capacity. We proposed following two CMA methods:

- User-antenna distance-based CMA,
- Centroid distance-based CMA.

In this this paper, after describing IC and CMA methods, we compare four UAC combinations comprehensively to find which combination is suitable for cluster-wise MU-MIMO system.

C. Orgnazition

The rest of this paper is organized as follows. Section II introduces the system model and channel model. We consider

978-1-7281-8296-4/20/$31.00 ©2020 IEEE

a cluster-wise zero-forcing (ZF)-based MU-MIMO. ZF is a well-known multi-user signal processing [6-8] to cancel the inter-user interference (IUI). In Section III, we propose UAC consisting of IC and CMA methods. Examples of cluster formation are illustrated. Then, simulation results on link capacity are presented in Section IV to compare possible combinations of two IC methods and two CMA methods. Then, the impacts of cluster number, transmit SNR, and propagation environment on the achievable link capacity are examined. Finally, some conclusions are given in Section V.

II. System Model

A. Downlink transmission model

As illustrated in Fig. 1, we consider the downlink communication of a cluster-wise ZF-based MU-MIMO system with ZF precoding. As a typical distributed antenna system, the BS is connected with a large number of distributed antennas deployed in different locations through optical mobile fronthaul. The BS communicate with U users via A distributed transmit antennas. All signal processing is carried out in the BS, and we assume that each user in the BS coverage is equipped one receive antenna. Then users and antennas are grouped into K clusters by proposed clustering and CMA methods based on the real distance. For each cluster, ZF precoding is carried out by assuming the BS gets perfect global channel state information (CSI), aiming to eliminate the IUI and enhance the system spectral efficiency.

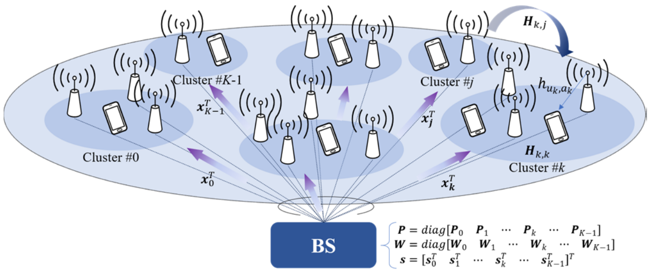


Fig. 1. Downlink transmission model of cluster-wise MU-MIMO.

Because pseudo-inverse matrix is used in ZF precoding for each cluster, we typically have $U \le A$ and $U_k \le A_k$. Here, U_k and A_k denotes the number of users and the number of antennas in the kth cluster with satisfying $\sum_{k=0}^{K-1} U_k = U$ and $\sum_{k=0}^{K-1} A_k = A$. Similarly, the u_kth user and the a_kth antenna in the kth cluster is presented as $u_k \in \{0_k, 1_k, \cdots, U_k - 1\}$ and $a_k \in \{0_k, 1_k, \cdots, A_k - 1\}$ respectively.

Mathematically, the totality precoded transmit signal from the antenna side can be represented as

$$\mathbf{x} = \mathbf{W}\sqrt{\mathbf{P}}\mathbf{s}, \tag{1}$$

where $\mathbf{P} \in \mathbb{R}_{\ge 0}{}^{U \times U}$ denotes the power allocation matrix, $\mathbf{W} \in \mathbb{C}^{A \times U}$ denotes the ZF precoder in BS, and $\mathbf{s} = \mathbb{C}^{U \times 1}$ denotes the transmit symbol vector satisfying $\mathrm{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_U$. As the power constraint, we assume that total transmit SNR in the kth cluster depends on U_k, i.e.,

$$\mathbf{P}_k = \frac{U_k \times P}{\|\mathbf{W}_k\|_F^2} \mathbf{I}_{U_k}, \tag{2}$$

where $\mathbf{P}_k \in \mathbb{R}_{\ge 0}{}^{A_k \times A_k} \subseteq \mathbf{P}$ denotes power allocation for the kth cluster, $\mathbf{W}_k \in \mathbb{C}^{A_k \times U_k} \subseteq \mathbf{W}$ denotes the precoder for the kth cluster, and P is the normalized transmit signal-to-noise ratio (SNR) we set for a single user. Here, $\|\bullet\|_F$ denotes the Frobenius norm, and $\mathbf{I}_N$ denotes the $N \times N$ identity matrix.

The communication channel $\boldsymbol{H}$ can be expressed in cluster-level as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 \\ \vdots \\ \mathbf{H}_{K-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{0,0} & \cdots & \mathbf{H}_{0,K-1} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{K-1,0} & \cdots & \mathbf{H}_{K-1,K-1} \end{bmatrix}, \tag{3}$$

where $\mathbf{H}_{k,j} \in \mathbb{C}^{U_k \times A_j} (k, j \in \{0, 1, \cdots, K-1\})$ denotes the channel between users in the kth cluster and antennas in the jth cluster. Assuming that the propagation channel is modelled as the composition of pathloss, log-normal shadowing loss, and fading, the channel gain $h_{u_k,a_k} \in \mathbf{H}$ can be expressed as

$$h_{u_k,a_k} = \sqrt{d_{u_k,a_k}{}^{-\alpha} 10^{-\frac{\varphi_{dB}}{10}}}\, z, \tag{4}$$

where d_{u_k,a_k} is the distance between the u_kth user and the a_kth antenna in the kth cluster, α is the pathloss exponent, φ_{dB} is a real-valued zero-mean Gaussian random variable with standard deviation of σ_{dB}, and z is a complex-valued zero-mean Gaussian random variable with unit variance which represents Rayleigh fading.

With the global channel state information (CSI), the ZF precoder which is carried out at BS for each cluster can be expressed as

$$\mathbf{W}_k = \mathbf{H}_{k,k}{}^{\dagger} = \mathbf{H}_{k,k}{}^{H} (\mathbf{H}_{k,k} \mathbf{H}_{k,k}{}^{H})^{-1}, \tag{5}$$

where $\boldsymbol{M}^{\dagger}$ denotes the pseudo-inverse of matrix $\boldsymbol{M}$.

Then let $\mathbf{R} = \mathbf{H}\mathbf{W}\sqrt{\mathbf{P}} \in \mathbb{C}^{U \times U}$, the received symbol at the u_kth user in the kth cluster can be given as

$$y_{u_k} = \mathbf{R}(u_k, :)\mathbf{s} + n_{u_k}, \tag{6}$$

where $\boldsymbol{M}(i,:)$ denotes the ith row of $\boldsymbol{M}$, and n_{u_k} denotes the complex Gaussian noise with mean 0 and covariance 1.

Furthermore, based on (6), the signal-to-interference-plus-noise-ratio (SINR) of the u_kth user in the kth cluster can be obtained as

$$SINR_{u_k} = \frac{|R(u_k, u_k)|^2}{\sum_{u_j=0, u_j \ne u_k}^{U-1} |R(u_k, u_j)|^2 + 1}, \tag{7}$$

where $\boldsymbol{M}(i,j)$ denotes the (i,j)th element of $\boldsymbol{M}$.

From this, the link capacity can also be computed by

$$\begin{cases} C_{u_k} = \log_2(1 + SINR_{u_k}) \\ C_{sum} = \sum_{k=0}^{K-1} \sum_{u=0}^{U_k-1} C_{u_k} \end{cases}. \tag{8}$$

B. Problem Formulation for Structure of Cluster-wise MU-MIMO

As the expression in (3) and (7), for the same distribution of users and antennas, different UAC results lead to different system capacity. According to this, how to form the cluster-wise MU-MIMO becomes the key issue, which is also our contribution. The criterion for UAC and which UAC approach is more suitable for this system is what we are concerned about.

III. USER-ANTENNA CLUATERING

In this section, we describe the proposed UAC consisting of IC and CMA. Firstly, we present two IC methods using K-means clustering algorithm based on user locations and antenna locations. Then, we present two CMA methods. There are possible four combinations of IC and CMA for performing cluster-wise MU-MIMO.

A. IC methods using K-means clustering algorithm

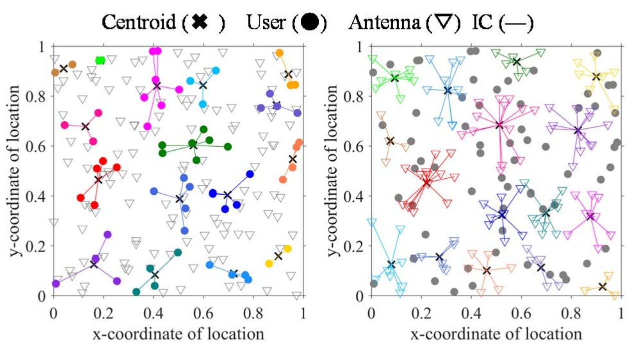
Aiming to reduce the huge computational complexity due to large-scale MU-MIMO, K-means clustering algorithm with a Euclidean metric for users and antennas is proposed. As a classic clustering method, K-means groups users or antennas into K clusters separately, such that each user or antenna belongs to the cluster whose centroid is the closest. The concrete description is given as Algorithm 1.

Algorithm 1 K-means clustering

1: Centroid initialization:
2: **repeat**
3: Choose one centroid uniformly at random among the user/antenna positions.
4: For each user/antenna, compute its distance to the nearest centroid that has already been chosen.
5: Choose one new user/antenna position at random as a new centroid, using a weighted probability distribution where user/antenna is chosen with probability proportional to the square of distance.
6: **until** K centroids have been chosen.
7: Standard K-means clustering:
8: **repeat**
9: For each user/antenna, compute its distance to all centroids and assign it to the closest cluster.
10: Update the cluster centroids based on the locations of their members.
11: **until** clusters no longer change.

Here we introduced k-means++ [5] algorithm for choosing the initial centroids to avoid the sometimes poor clustering results found by the standard K-means.

For simplicity, in this paper, the BS coverage is normalized as a 1 by 1 square, where users and distributed antennas are randomly uniform distributed. Specifically, as for a certain cluster-wise MU-MIMO system containing 128 antennas and 64 users to meet the restriction of ZF precoding, the 16-cluster clustering for users and antennas are illustrated in Fig. 2 separately.



(a) User-clustering (b) Antenna-clustering

Fig. 2. K-means clustering for a certain user-antenna distribution. A=128, U=64, K=16.

Fig. 2 shows that K-means algorithm can well group users and antennas according to the number of clusters given in advance, and get good separated clusters. With the allocation of the first process, antennas and users need to be assigned to the corresponding clusters to form the complete cluster-wise MU-MIMO. We considered the location relationship between users and antennas in the BS coverage to assign them by CMA methods in III.B and III.C.

B. Balanced user-antenna distance-based CMA

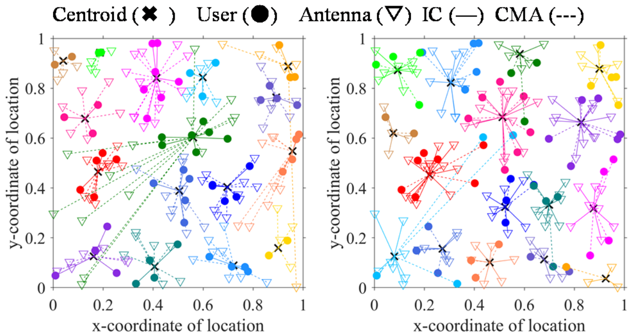
Shortening the propagation path can effectively reduce propagation loss. According to this, we first considered the shortest distances between users and antennas to assign users to antenna-clusters or to assign antennas to user-clusters. Moreover, because the K-means algorithm is spontaneous convergence, the number of members in each cluster is determined by their locations. So, we want to keep a global balance in this process of allocating members, that is, to keep the proportion of the total number of antennas and the total number of users in each cluster as far as possible. Combining these two points, balanced user-antenna (U-A) distance-based CMA is described as Algorithm 2.

Algorithm 2 Balanced U-A distance-based CMA

1: Initialization: Calculate distance between each user and antenna, then sort the distances in ascending order.
2: **Case 1**: user-clustering (UC)
3: Step 1: Assign antennas to user clusters which their closest user belongs to. Iterate in ascending order of U-A distances, until all user clusters satisfy $A_k = \lfloor U_k \times A \div U \rfloor$.
4: Step 2: Assign remaining antennas to user clusters which their closest user belongs to. Iterate in ascending order of rest U-A distances, until all antennas are assigned while $\forall U_k \le A_k$.
5: **Case 2**: antenna-clustering (AC)
6: Step 1: Assign users to antenna clusters which their closest antenna belongs to. Iterate in ascending order of U-A distances, until all antenna clusters satisfy $U_k = \max(\lfloor A_k \times U \div A \rfloor, 1)$.
7: Step 2: Assign remaining users to antenna clusters which their closest antenna belongs to. Iterate in ascending order of rest U-A distances, until all users are assigned while $\forall U_k \le A_k$.

Both for user clustering and antenna clustering, two-step process CMA is formed. Their same principle is to ensure that the number of users and antennas in each cluster meet the restriction of utilizing ZF. Then main difference between the two cases is Step 1. We introduced the totality user-to-antenna ratio to control the equalization of this ratio in each cluster. Here, $\lfloor n \rfloor$ denotes the greatest integer less than or equal to n. It is noteworthy that for the case of antenna clustering, if the number of users in the system is small and the number of antennas in a certain cluster is small (extreme case is 1), then the result of the floor function may become 0, so we attach a term to ensure that the number of users in the antenna cluster is at least one.

Based on the clustering result shown in Fig. 2, the final cluster containing users and antennas formed by Algorithm 2 is shown as Fig. 3. Fig. 3 shows that both the antennas and users are assigned to the corresponding clusters, the more users in cluster, the more antennas there are, and vice versa. Nevertheless, selection of remote antennas or users occurs, which may affect the system capacity.



(a) User-clustering (b) Antenna-clustering

Fig. 3. Cluster assignment by balanced U-A distance-based CMA. A=128, U=64, K=16.

C. *Centroid distance-based CMA*

In order to avoid the remote selections in Algorithm 2, we attempt to design CMA from another perspective. Inspired by clustering, we consider each cluster as a whole to select antenna or user members to make the final clusters more compact. In other words, this time CMA is not based on U-A distance, but the distance from users or antenna to cluster centroids. Moreover, the restriction on the number of users and antennas in Algorithm 2 (step 1) is the main factor for the remote selection. Therefore, the restrictions on the number of users and antennas in each cluster are further relaxed to ensure that cluster-wise ZF precoding can be realized. The detailed description of centroid-based CMA is given as Algorithm 3.

Algorithm 3 Centroid distance-based CMA

1: **Case 1**: user-clustering (UC)

2: Initialization: Do user-clustering as Algorithm 1. Calculate distance between each antenna and cluster centroid, then sort the distances in ascending order.

3: Step 1: Assign antennas to user clusters which their closest centroid belongs to. Iterate in ascending order of A-C distances, until all user clusters satisfy $A_k = U_k$.

4: Step 2: Assign remaining antennas to user clusters which their closest centroid belongs to. Iterate in ascending order of rest A-C distances, until all antennas are assigned while $\forall U_k \leq A_k$.

5: **Case 2**: antenna-clustering (AC)

6: Initialization: Do antenna-clustering as Algorithm 1. Calculate distance between each user and cluster centroid, then sort the distances in ascending order.

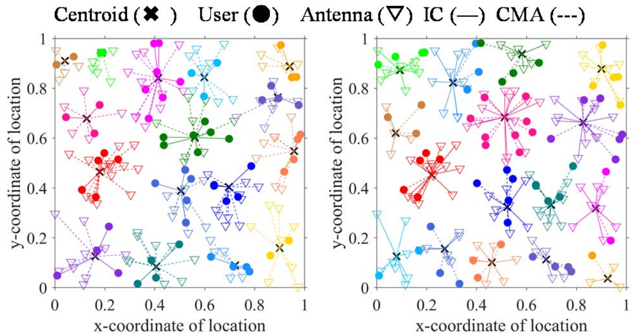
7: Step 1: Assign users to antenna clusters which their closest centroid belongs to. Iterate in ascending order of U-C distances, until all user clusters satisfy $U_k = 1$.

8: Step 2: Assign remaining users to antenna clusters which their closest centroid belongs to. Iterate in ascending order of rest U-C distances, until all antennas are assigned while $\forall U_k \leq A_k$.

Compared with U-A distance-based CMA, this centroid-based CMA also maintains the two-step allocation. The restrictions are reduced to the lowest level, and for the second step, they are unchanged. In addition, the U-A distance set can be calculated once before clustering, but centroid locations are needed to calculate the distance, which means clustering needs to be done at the beginning. As a contrast, the result of Fig. 2 clusters attached centroid-based CMA is plotted in Fig. 4.

It can be clearly seen that the final clusters after user or antenna assignment become more compact. In the case where the number of users is smaller than the number of antennas in the example, remote selection can be well eliminated.

So far, both CMA methods have been introduced, with different emphasis on their algorithm design, which also leads to dissimilar in graphic impressions. As the construction of communication system, which CMA is better should be evaluated by link capacity.



(a) User-clustering (b) Antenna-clustering

Fig. 4. Cluster assignment by centroid distance-based CMA. A=128, U=64, K=16.

IV. Link Capacity Evaluation

In this section, link capacity of cluster-wise MU-MIMO system formed by UAC is evaluated. 4 combinations of IC methods and CMA methods are compared through Monte Carlo simulation of cumulative distribution function (CDF) of the sum capacity. Then, we comprehensively discuss the impacts of the number of clusters and various propagation environments on the link capacity and also a trade-off between link capacity and computational complexity.

A. *Simulation setting*

Simulation settings are shown in Table I. Throughout the simulation, we use a fixed pattern of 128 antennas whose locations are randomly generated. For the given antenna location pattern, user locations are randomly generated 1,000 times to calculate the link capacity to obtain its CDF.

TABLE I. Simulation Settings

Parameter	Value/State
No. of distributed antennas	128
No. of users	16, 32, 48, 64, 80, 96, 112, 128
No. of clusters	1, 2, 4, 8, 16, 32, 64
Pathloss exponent	3.5
Shadowing standard deviation	8 dB
Fading type	Frequency-nonselective Rayleigh fading
Normalized transmit SNR	-30 dB, -20 dB, -10 dB, 0 dB
Number of times of user location change	1000
Number of times of shadowing change	10
Number of times of fading change	10

The normalized transmit signal-to-noise ratio (SNR) of x dB is defined as the transmit SNR which achieves the received SNR of x dB when the distance between the transmitter and receiver is equal to the side length of square-shaped BS area when shadowing and fading do not exist. Every time we change the user distribution, we carry out UAC to form clusters again and compute the pathloss of the propagation channel between antenna and user. For each user distribution, the log-normal shadowing losses are changed 10 times. The Rayleigh fading gains are changed for each set of shadowing losses to compute the local average sum capacity (i.e., the sum capacity averaged over Rayleigh fading for the given path loss and shadowing loss).

B. Simulation Results

1) Comparison among 4 UAC combinations

Here we plot the 50% value of sum capacity CDF curve of different number of users with the same 8 clusters to compare 4 combinations of cluster assignment in Fig. 5.

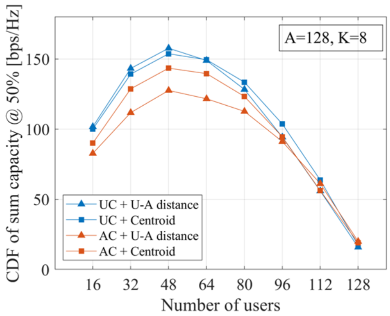


Fig. 5. Comparison of 4 UAC combinations in 50% CDF of sum capacity among different number of users with -20 dB transmit SNR.

As can be seen in Fig. 5, increasing the number of users will reduced the sum capacity of system. Because we only have ZF precoding in each cluster, the inter-cluster interference (ICI) still exists, increasing of the number of users will reduce the number of users and antennas in a single cluster, and make more users affected by ICI. When the number of users is far less than the number of antennas, the capacity is not very high, because small-scale MIMO is difficult to make full use of the diversity and multiplexing gains.

Then, four ways of UAC are compared vertically. First, the change trend of the four curves is basically the same among different users. In general, user clustering achieves higher capacity except for the same number of users and antennas. In our system scenario, the users' locations change while the antennas' locations are fixed. Therefore, the clustering method which can better adapt to changing users' locations is the UC. Then, in the case of user clustering, two types of CMA perform nearly the same. On the other hand, in the case of antenna clustering, the centroid-based CMA performs better mostly. Combining these, user clustering with centroid-based CMA is preferred normally.

Besides, the results plotted in Fig. 5 are based on the same antenna distribution as in Fig. 2. Without loss of generality, we also do the same thing with several different antenna distributions and different K, although the specific values are high or low, then trend of four curves is the same as that shown in Fig. 5. That means under a large number of user distribution samples, a specific antenna distribution has no effect on the comparison results.

2) Impacts of other parameters

In this part, we choose the UAC with user-clustering and centroid-based CMA to find out impacts of several other parameters on the performance of cluster-wise MU-MIMO.

First, we compare the link capacities achievable in different propagation environments. It can be seen from Fig. 6 that the worst environment is the case of path loss only and the best environment is the case with shadowing and fading. This is because spatial diversity gains and multi-user diversity gains can be obtained thanks to spatially distributed antennas and users.

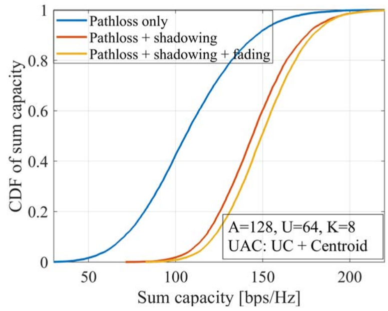


Fig. 6. Impact of propagation environment on sum capacity.

Then the cumulative distribution of the sum capacity is plotted in Fig. 7 for different normalized transmit SNRs. It can be seen that the sum capacity is kept almost the same even if the transmit SNR is reduced to -20dB. This suggests that clustering can be effective to reduce the transmit SNR, thereby improving the energy efficiency. This is because of distributed nature of antennas and users; almost always each user can find at least one antenna which is very close to it.

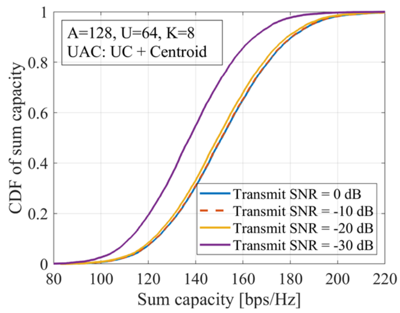


Fig. 7. Impact of transmission SNR on sum capacity.

The sum capacity and the computational complexity are plotted as a function of the number of clusters in Fig. 8. Here, the computational complexity is determined by the number of floating-point operations required for performing the ZF. ZF requires the matrix inversion whose complexity is approximately given as $O(u_k^3)$ [7].

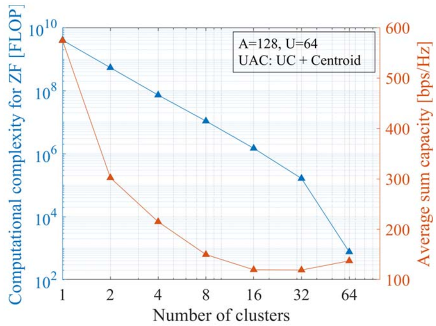


Fig. 8. Impact of number of clusters on sum capacity and computational complexity.

It can be seen that both the sum capacity and the computational complexity decrease with the increase of the number of clusters especially when the number of clusters increases from one (no clustering) to two and then the decreasing rate trends to become gentler. It should be noticed that the computational complexity reduces much more steeply than the sum capacity. As an extreme case when the number of clusters is equal to the number of users, cluster-wise MU-MIMO becomes multi-input single-output (MISO). If the number of clusters is less than 10, the sum capacity is better kept above the MISO situation, otherwise (more than 10 clusters) like the MISO situation.

In Fig. 8, we want to point out that $K = 1$ has the largest capacity, but its high computational complexity makes it impractical to implement. And the sum capacity decreases with the increase of K, which is caused by increasing ICI. The ICI can be reduced by using some resource allocation algorithm, in which the available frequency or time resource is divided into several orthogoal blocks and they are allocated to different clusters [13].

V. Conclusion

In this paper, we considered ZF-based MU-MIMO in an ultra-dense 5G advanced systems. Trying to reduce the high computational complexity caused by the large-scale MU-MIMO to reduce the complexity to an acceptable level, we introduced clustering to decompose the large-scale MU-MIMO into smaller-scale ones. We proposed the user-antenna clustering (UAC), which consists of the initial clustering (IC) using K-means algorithm to form either user clusters or antenna clusters and the cluster member assignment (CMA) to allocated users or antennas into the initial clusters.

We compared four possible combinations of two IC methods and two CMA methods through Monte Carlo simulation of the achievable sum capacity with cluster-wise ZF-based MU-MIMO. It was shown that user clustering can achieve a higher capacity under various conditions of number of users and that of clusters for the given number of distributed antennas. Performance of two CMA methods are similar, but the centroid-based one is a preferable choice from the computational complexity point of view. Furthermore, we comprehensively discussed other parameters in cluster-wise MU-MIMO. It was found that multipath-rich propagation environment is beneficial to distributed cluster-wise MU-MIMO. Confirm to our expect, clustering can effectively reduce computational complexity and save transmit SNR. And for a certain number of antennas, the number of users and that of clusters have a great impact on the achievable link capacity. Optimized number of users and clusters with trade-off between link capacity and computational complexity is left as our future study.

In this paper, the same transmit SNR was assumed for all users. However, a different user suffers from a different inter-cluster interference, Therefore, the power allocation among the users to maximize the sum capacity or user capacity while keeping the total transmit SNR the same is an important issue for our future study. We focused on a single-cell system in this paper. But in multi-cell systems, the inter-cell-interference is present and limits the capacity. The performance of clustering method in a multi-cell system is left as our future study.

Acknowledgment

A part of this work was conducted under “R&D for further advancement of the 5th generation mobile communication system” (JPJ000254) commissioned by the Ministry of Internal Affairs and Communications in Japan.

References

[1] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, “Cellular architecture and key technologies for 5G wireless communication networks,” IEEE Communications Magazine, Vol. 52, Issue 2, pp. 122-130, Feb. 2014.

[2] W. Roh, J. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, “Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results,” IEEE Commun. Mag., Vol. 52, Issue 2, pp. 106-113, Feb. 2014.

[3] J. Joung, Y. K. Chia, and S. Sun, “Energy-efficient, large-scale distributed-antenna system (L-DAS) for multiple users,” IEEE J. Selected Topics in Signal Processing, Vol. 8, No. 5, pp.954-965, Oct. 2014.

[4] J. A. Hartigan and M. A. Wong, “Algorithm AS 136: A k-Means clustering algorithm,” J. Royal Statistical Society, Series C (Applied Statistics), Vol. 28, No. 1, pp. 100-108, 1979.

[5] D. Arthur and S. Vassilvitskii, “k-means++: the advantages of careful seeding,” Proc. 18th Annual ACM-SIAM Symposium on Discrete Algorithms, New Orleans, Louisiana, USA, 7-9 Jan. 2007.

[6] A. Wiesel, Y. C. Eldar, and S. Shamai, “Zero-forcing precoding and generalized inverses, ” IEEE Transactions on Signal Processing, vol. 56, no. 9, pp. 4409-4418, Sept. 2008, doi: 10.1109/TSP.2008.924638.

[7] C. Park, Y. Byun, A. M. Bokiye, and Y. Lee, “Complexity reduced zero-forcing beamforming in massive MIMO systems, ” Proc. 2014 Information Theory and Applications Workshop (ITA), San Diego, CA, 2014, pp. 1-5, doi: 10.1109/ITA.2014.6804219.

[8] C. Zhang, Y. Jing, Y. Huang, and L. Yang, “Performance analysis for massive MIMO downlink with low complexity approximate zero-forcing precoding,” IEEE Trans. Commun., Vol. 66, Issue 9, pp.3848-3864, Sept. 2018.

[9] L. Zhang, L. Gui, K. Ying, and Q. Qin, “Clustering based hybrid precoding design for multi-user massive MIMO systems, ” IEEE Transactions on Vehicular Technology, vol. 68, no. 12, pp. 12164-12178, Dec. 2019, doi: 10.1109/TVT.2019.2949578.

[10] T. Kim, “User scheduling and grouping in massive MIMO broadcast channels with heterogeneous users, ” Journal of Communications and Networks, vol. 21, no. 4, pp. 385-394, Aug. 2019, doi: 10.1109/JCN.2019.000028.

[11] K. Hosseini, W. Yu, and R. S. Adve, “Cluster based coordinated beamforming and power allocation for MIMO heterogeneous networks, ” Proc. 2013 13th Canadian Workshop on Information Theory, Toronto, ON, 2013, pp. 96-101, doi: 10.1109/CWIT.2013.6621600.

[12] Z. Jiang, S. Zhou, and Z. Niu, “Optimal antenna cluster size in cell-free large-scale distributed antenna systems with imperfect CSI and intercluster interference,” IEEE Transactions on Vehicular Technology, vol. 64, no. 7, pp. 2834-2845, July 2015, doi: 10.1109/TVT.2014.2349959.

[13] C. Ge, S. Xia, Q. Chen, F. Adachi, “2-step graph coloring algorithm for cluster-wise distributed MU-MIMO in ultra-dense RAN,” to be presented at The 23rd International Symposium on Wireless Personal Multimedia Communications (WPMC 2020), Virtual Conference, 19 - 26 Oct. 2020.